

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing in the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any of reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

AFRL-SR-BL-TR-00-

0169

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 3/7/2000		3. REPORT TYPE AND D Final	
4. TITLE AND SUBTITLE Spatial Soliton Interactions for Photonic Switching Part II				5. FUNDING NUMBERS AASERT F49620-95-1-0432	
6. AUTHOR(S) Kelvin H. Wagner and Steve Blair					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Colorado at Boulder Optoelectronic Computing Systems Center Campus Box 525 Boulder, CO 80309				8. PERFORMING ORGANIZATION REPORT NUMBER N/A	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Dr Alan Craig AFSOR, Bldg 410 110 Duncan Ave, Suite B115 Bolling AFB, Washington DC 70332				10. SPONSORING / MONITORING AGENCY REPORT NUMBER 20000503 096	
11. SUPPLEMENTARY NOTES Companion report is F49620-95-1-0431					
12a. DISTRIBUTION / AVAILABILITY STATEMENT Full version of both reports available online at http://optics.colorado.edu/~kelvin/links/soliton_report.ps.gz				12b. DISTRIBUTION CODE UL	
13. ABSTRACT (Maximum 200 Words) High-bandwidth optical communications will greatly benefit from optical switches since they could eliminate the optical/electronic conversion. Optical logic gates allowing data regeneration, gain, cascability, would allow even more complex all-optical routing functions. In this work we report on an in-depth study of an optical logic gates based on spatial and spatio-temporal solitons. Optical solitons that propagate long distances without change, act as the natural carrier of binary data due to their stability to perturbations and intrinsic threshold. The non-diffracting nature of spatial optical solitons lends to their use in a class of angular deflection logic gates, in which a weak signal can alter the propagation of a strong pump in order to change the device state from high to low, thereby implementing a controlled inverter which is cascable to produce logically-complete, multi-input NOR. This work develops theoretical and numerical framework to describe general, multi-dimensional, nonlinear spatio-temporal wave phenomena. From Maxwell's equations we derive via the multiple-scales perturbation technique a first-order, fully-vectorial, nonlinear wave equation, that is valid beyond the standard slowly-varying amplitude, slowly-varying envelope, and paraxial approximations. In addition to coupling with the orthogonal transverse field, vector coupling with the weak longitudinally-projected field is also treated, along with the cascaded interaction with a weak third-harmonic wave which can produce a desirable saturation effect.					
14. SUBJECT TERMS Photonic switching, optical logic, nonlinear optics, optical solitons, spatiotemporal solitons, multiple scales, derivation, nonlinear wave equation				15. NUMBER OF PAGES	
				16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL		

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

DTIC QUALITY INSPECTED 3

Part II

AASERT Final report F49620-95-1-0432

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

II AASERT Final report F49620-95-1-0432

1	Introduction, Motivation and Background	4
1.1	Outline	4
1.2	Applications in Switching and Computing	5
1.2.1	Optical Communications Networks	5
1.2.2	Digital Optical Computing and Processing	6
1.3	Requirements for Digital Switching and Logic Devices	7
1.3.1	Switching Devices	7
1.3.2	Logic Devices	7
1.3.3	Optical Devices	9
1.4	Digital Optical Switching and Logic Devices	9
1.4.1	Nonlinear Directional Coupler	9
1.4.2	Kerr Gate	13
1.4.3	Nonlinear Optical Loop Mirror	18
1.4.4	Terahertz Optical Asymmetric Demultiplexor	20
1.5	Optical Soliton Logic Devices	23
1.5.1	Temporal Soliton Dragging Gate	24
1.5.2	Angular Deflection Gate based on Spatial Soliton Dragging	27
2	Optical Solitons	33
2.1	Historical Background	33
2.2	Preliminaries	35
2.3	1-D Spatial Optical Solitons	35
2.3.1	The Non-Paraxial Fundamental Soliton	37
2.3.2	The Paraxial Fundamental Soliton	37
2.4	1-D Temporal Optical Solitons	42
2.4.1	The Non-SVEA Fundamental Soliton	43
2.4.2	The SVEA Fundamental Soliton	44
2.4.3	Higher-Order 1-D Temporal NLS Solitons	48
2.4.4	Higher-Order Temporal Effects	49
2.5	2-D and 3-D Spatio-Temporal Solitary Waves	53
5	1-D Spatial Soliton Logic Gates	60
5.1	Soliton Interactions	61
5.1.1	Definition of Parameters	62
5.1.2	Phase-Sensitive Interactions	64
5.1.3	Phase-Insensitive Interactions	72
5.2	Soliton Collision and Dragging Logic Gates	79
5.2.1	The Soliton Collision Logic Gate	81
5.2.2	The Soliton Dragging Logic Gate	83
5.2.3	Comparison of Logic Gates	84
5.3	Effects of Material Absorption	86
5.3.1	Adiabatic Evolution of a Single Soliton	89
5.3.2	Figures-of-Merit for Example Materials	94
5.3.3	Comparison of Logic Gates with Absorption	95
5.4	Cascading of Collision and Dragging Gates	98
5.4.1	Gate Transfer Function	101
5.4.2	Ring Oscillator Configuration	106
5.4.3	Two-Input NOR Configuration	108

6	2-D Spatio-Temporal Solitary Wave Logic Gates	118
6.1	(2+1)-D Spatio-Temporal Propagation and Stability	118
6.1.1	Propagation without Saturation	120
6.1.2	Propagation with Saturation	124
6.2	2-D Spatio-Temporal Dragging Logic Gates	128
6.3	Cascaded Spatio-Temporal Dragging Logic Gates	138
6.3.1	Ring Oscillator Configuration	139
6.3.2	Two-Input NOR Configuration	146
6.4	Effects of Material Properties	148
7	Conclusions	150

Chapter 1

Introduction, Motivation and Background

This thesis deals with the topic of ultrafast, all-optical switching and logic using spatial and spatio-temporal solitons. The main focus is on the accurate analytical and numerical modeling of the interaction among solitons in geometries that allow for logically-complete, cascaded logic gates with fanout and level restoration. Detailed modeling will assist in subsequent experimental investigation by identifying regions of stability, robust operation and reliability, and candidate material systems.

The work in this thesis makes significant contributions in the areas of fundamental nonlinear spatio-temporal propagation and optical switching and logic. The first contribution is the result of a multiple-scales derivation directly from Maxwell's equations, which results in a first-order, multi-dimensional, vectorial differential equation that is accurate beyond the standard paraxial and slowly-varying envelope and amplitude approximations. This result is fundamental to the studies of this thesis and will have application in other areas as well, such as optical communications, short-pulse passively mode-locked lasers, and atmospheric pulse propagation.

The second contribution arises from the systems-level approach taken to optical switching and logic. Most optical switching technologies do not satisfy the basic requirements for a logic gate, thereby rendering them of little use in applications beyond simple, single-stage switching operations. The novel logic gates studied here satisfy the necessary requirements and thus have the potential to make an impact in areas of ultra-high speed switching and logic systems. The desirable properties of these logic gates are facilitated by the non-diffracting and/or non-dispersing nature of optical solitons.

1.1 Outline

The remainder of this chapter briefly presents background in device requirements and optical switching and logic. Section 1.2 provides motivation and a brief overview of promising applications for optical logic. Section 1.3 covers generic requirements for switching and logic devices and additional specific requirements for optical devices. A brief review of other contemporary optical switching devices and their shortcomings is presented in section 1.4. Logic devices which specifically take advantage of the properties of optical solitons are discussed in section 1.5, which also provides an introduction to the optical soliton logic gates studied in this thesis.

Detailed background on optical solitons is given in Chapter 2. First, section 2.1 provides historical background in solitary wave and soliton phenomena. Section 2.2 covers some preliminary details that lead directly to the discussion of optical solitons and solitary waves. The following sections then discuss spatial (sec. 2.3), temporal (sec. 2.4), and spatio-temporal (sec. 2.5) optical solitons.

Chapter 3 derives the multi-dimensional vector wave equations necessary for the numerical simulations of the soliton-based logic gates presented in later chapters. Section 3.1 covers a standard treatment of a fully vectorial, nonlinear Helmholtz-type wave equation along with background on the linear and nonlinear susceptibilities. This second-order wave equation is unsuitable for the purposes of this thesis due to the difficulty in numerically propagating an initial field distribution because of the fast time and space scales in the equation. Instead, section 3.2 derives an asymptotic vector nonlinear wave equation directly from Maxwell's equations which is first-order in the propagation coordinate and depends only on the scales of the slowly-varying envelope, and is thus more suitable for numerical propagation. The importance of the results obtained in this chapter arises from the higher-order nature of the derived equation, which extends beyond the approximations made in the well-known multi-dimensional nonlinear Schrödinger (NLS) equation, resulting in the capability to describe propagation of vector optical field envelopes with very broad spectral content in both the spatial and temporal frequency domains, including the effects of higher-order nonlinearities.

Chapter 4 describes the split-step numerical method used to solve the vector nonlinear wave equations. The basis of

this method is to treat linear and nonlinear (inhomogeneous) propagation in separate steps, performing linear propagation in its natural Fourier domain and nonlinear propagation in the real-space domain. For a small step size, this splitting is a good approximation. Section 4.1 covers multi-dimensional linear spatio-temporal diffraction, which can be solved without approximation as an initial value problem. Noting that linear and nonlinear propagation are not separable, section 4.2 derives a split-step formulation which is approximately second-order accurate in step size. This formulation is then applied separately to (1+1)-D spatial and (2+1)-D spatio-temporal nonlinear propagation, to be used in later chapters. Finally, section 4.3 discusses the accuracy of the method as it applies to problems of later chapters.

The use of one-dimensional spatial soliton interactions for logic gates is presented in Chapter 5. The first section covers the basic soliton interaction geometries useful for three-terminal, restoring logic, noting that the collision and dragging geometries using orthogonally polarized solitons, which are of the general class of angular deflection gates, provide the best performance in terms of large gain with high contrast. Section 5.3 then examines the effects of linear and two-photon absorption on the propagation of a single spatial soliton and develops appropriate figures-of-merit for use in comparing the suitability of nonlinear materials for soliton logic gates. The collision and dragging logic gates are then compared in the presence of absorption using the material parameters of fused silica, where the spatial dragging gates generally perform better because of their shorter gate lengths. The final section (5.4) computes the transfer functions of the collision and dragging gates and determines the maximum fanout that these gates can provide in a cascaded system. The results of this chapter provide guidance for the spatio-temporal logic gates of the next chapter, which is of ultimate interest.

Logic gates based on the interactions between two-dimensional spatio-temporal solitary waves is the subject of Chapter 6, which have the potential for greater than THz switching rates with nJ to pJ switching energies. Section 6.1 discusses the propagation of a single spatio-temporal solitary wave with higher-order linear and nonlinear effects as derived in Chapter 4. This section shows numerically that stabilized solitary waves suitable for logic gates should exist. With these results, section 6.2 studies the vectorial interaction between these spatio-temporal solitary waves with emphasis on the dragging interaction. For completeness, section 6.3 presents results on cascaded logic. It is shown that the dragging logic gates can provide reasonable fanout in a cascaded, logically-complete, system, but also that, ultimately, the performance of the logic gate will be limited by the effects of Raman scattering.

Finally, Chapter 7 concludes, noting that this thesis provides the theoretical and numerical development that is necessary to study a novel class of all-optical logic gates. Numerical simulations have shown that these gates satisfy the system requirements that suggest their use beyond simple, single-stage operation, and paves the way for future theoretical and experimental work on ultrafast, all-optical logic systems.

1.2 Applications in Switching and Computing

The traditional advantages that optics hold over electronics are the inherent speed, parallelism, and lack of inductive or capacitive crosstalk, although crosstalk does occur to some extent in linear propagation due to scattering and diffraction, and in nonlinear propagation, which plays an important role in the topic of this thesis. These properties allow linear optics to map well onto problems in interconnection [1]. It is generally accepted that optical interconnects will replace electric transmission lines in many applications. Indeed, this is already occurring in long-haul communications networks and is expected to occur in short-haul and local-area networks as well. It is an open question to determine how far down the interconnect hierarchy this trend will reach, but it is fairly clear that it will extend to system level switching fabrics. In all of these cases though, the data is actually switched electronically, or under electronic control, and processed electronically.

It is less clear what role optics will play in switching and digital computing however. The soliton-based logic devices studied in this thesis are one approach to all-optical logic for switching and computing applications. Numerous all-optical switching and logic devices have been proposed, but none so far have reached the stage of practicality for real systems. There are certain minimum requirements for a logic device, and many of the optical approaches to logic fail to meet them. These requirements are discussed in section 1.3 and some contemporary optical switching devices and their failings are discussed in section 1.4. A clear understanding of these failings aids the present study of optical soliton-based logic devices. As section 1.5 explains, the soliton devices completely satisfy the fundamental requirements (and most of the practical requirements as well) for digital logic [2,3] and might succeed where the other optical approaches have failed.

The following sections discuss two general areas for which restoring, cascable logic gates, such as those studied in later chapters of this thesis, might find widespread application.

1.2.1 Optical Communications Networks

With the advent of high-bandwidth optical communications [4], high-speed switching technology becomes a necessity. Because information will be transmitted optically, it makes sense to explore optical technologies and devices to perform (or at least assist

in) the switching necessary to multiplex, demultiplex, and route information to the correct destination. Future bandwidth needs will be too great for an electronic-only solution to the switching problem. An advantage of optical devices in communications systems is the ability to switch and/or perform logic at the bit rate. This means that switching and logic operations are performed at and scale with the rate of incoming data, which is especially important in Tb/s communications switching and routing applications. Conversely, interconnected electronic gates cannot arbitrarily scale with the data rates and may ultimately be limited to speeds of about 50 Gbit/s [5].

Optical switching and logic devices are well suited for time-division multiplexed (TDM) data transmission [6], in which many low bandwidth channels are interleaved into individual time slots within a single high bandwidth channel. Here, the aggregate data rate may exceed 100 Gb/s [7], which is well beyond the performance expectation of practical, low power, electronic or optoelectronic switching networks. Therefore, at the very least, optical switching technology will find use in the intermediate processing layer in which the individual channels are optically separated (demultiplexed) from the single, high-speed transmission channel, for subsequent processing in the electronic/optoelectronic domain. An additional area of use is in data regeneration within an optical transmission line, which requires an optical logic gate. Given sufficient noise margins, an optical logic gate regenerates a noisy or attenuated data stream with fresh pulses that are derived from a power supply or clock, such that timing, amplitude, shape, etc., are completely restored. This type of all-optical repeater may eliminate the need for high-speed, high-power electronic repeaters in long-haul transmission.

Currently, the most popular multiplexing technique for long-haul communications and local-area optical networks is wavelength division multiplexing (WDM), in which each low bandwidth source channel is assigned a slot in the optical frequency spectrum. WDM has the advantage that multiplexing and demultiplexing can be performed by spectrally selective, passive, optical elements. However, due to the difficulties in developing a large array of stable wavelength sources, the individual channels in a WDM system may exceed data rates of 10 Gb/s [8,9], which may be of sufficient bandwidths that other advantages of optical switching and logic devices, such as the elimination of the optoelectronic conversion process [10], potentially lower power, and the use of deep optical circulating delay lines [11] for contention resolution, may play a role in the choice of implementation technology.

More complex operations are required in packet switched optical networks [10,12]. In these networks, at each node, a packet header is decoded, which determines the direction in which to route the packet. Header recognition is a simple digital correlation operation, but must be performed at the bit rate. Typical implementations based on optical switches [13] perform this operation in parallel, which results in a $1/N$ loss where N is the number of bits in the header. In order to reduce the latency at each node and avoid the fanout loss, a digital comparator, or shift register, could be used instead, based on optical logic gates. In addition, optical logic could be used for contention resolution, real-time encryption/decryption, and other complex tasks as well. Therefore, intelligent digital optical processing may play a significant role in the development in the next generation of high-speed optical communications networks.

1.2.2 Digital Optical Computing and Processing

Determining what role optics should play in general-purpose computing is an open question. The technological lead, resources, and infrastructures that electronics enjoy may be insurmountable for general-purpose optical computing. Even a compelling optical technology may not be enough to impact the future of digital computing, but there is potential in niche areas which are discussed in this section.

Simple digital optical logic circuits have been demonstrated [14,15] and proposed [16], which pave the way for more complex systems. However, many of these studies have been based on optical switches or gates that do not completely satisfy the requirements for digital logic. For example, the nonlinear Fabry-Perot etalon [17] is a two-terminal device which must be biased about an operating point with a holding beam and is very sensitive to variations in device parameter (i.e. the transfer characteristics vary from device to device). All of these factors preclude their use in complex systems. The logic gates studied in this thesis, on the other hand, do not suffer from these problems and could serve as fundamental building blocks in more complex systems.

In the absence of high space-bandwidth optical interconnections, bit-serial computation and signal processing [18] is a promising application area, especially when the problems of interest can be processed in a highly pipelined manner [19]. General purpose optical computers [20] have even been demonstrated using the bit-serial approach. One potential limitation for ultra-fast digital optical computing is the memory store. It has been shown, however, that an optical delay line can be used to implement a general machine [16], at the cost of increased latency. These tradeoffs are based on the transformations of computational origami [21]. Even more computational capacity can be realized with the combination of optical logic gates with optical interconnection, for which the power of the interconnection network allows for the efficient mapping of problems that would be difficult to process electronically [22].

1.3 Requirements for Digital Switching and Logic Devices

Routing switches and logic elements must possess certain properties in order to function properly within a system. The major distinction made here between switching elements and logic gates is that switching elements physically direct electrons or photons from one or more input ports to one or more output ports, while logic gates replace electrons or photons on the input ports with "new" electrons or photons derived from a power supply which then appear at the output ports. The requirements for a switching device are much less stringent than those for a logic device and will be discussed first. Additional requirements for an optical logic device are discussed in the following section.

1.3.1 Switching Devices

Optical routing switches are typically used just in the first switching stage to reduce data bandwidth to a level that electronics can handle in subsequent stages. These switching elements use a control to physically direct light from one or more input ports to one or more output ports and thus can be used to implement multiplexing and demultiplexing functionality. The routing decision is based either on the intensity of the signal inputs or by the presence of an externally derived control which is independent of the switching fabric. Therefore, the control is typically not of the same format as the data, meaning that the output of the switch cannot directly serve as the control for another switch [3]. As a result, this type of routing switch has fewer requirements than a logic gate.

The most important requirement of a switching device is high contrast operation. In binary transmission, the contrast determines the threshold level which separates the high and low states. Higher contrast gives larger noise margin, resulting in lower bit-error-rate (BER). For the output of a switching element, the overall contrast depends both on the contrast of the data stream (the difference in signal level between the two binary states which depends on the modulation ratio at the input of the transmission line and on the transmission line itself) and the contrast of the switching operation.

Another important requirement is switch transparency. If the switch is lossy or only a small fraction of the input is diverted to the output by the switching action, then high-speed, single-shot detection may become impossible, resulting in information loss. This does not mean that the switching operation must provide gain, though. For high-contrast output, saturation of an external amplification stage can be used to restore the data signal levels, but the gain recovery time may limit the data rate and introduce inter-symbol interference, the contrast at the output of the amplifier may be reduced, and complete logic level restoration (as discussed later for optical devices) is not obtained. The amount of gain is also limited due to amplified-spontaneous emission (ASE) noise, which increases the noise floor.

Routing switches do not have signal level restoration (in the absence of external amplification or level shifting) and the outputs may degrade due to loss or crosstalk and therefore BER may suffer from a long cascade of switches. If data pass through many levels of switching, such as in a multi-level implementation of an $N:N$ crossbar or the binary tree structure required for $1:N$ or $N:1$ multiplexors or demultiplexors constructed from $1:2$ or $2:1$ switches, then gain and level restoration as mentioned previously may be a necessity, subject to the limitations discussed in addition to ASE noise. Three-terminal operation also becomes necessary when many levels of switching elements are used such that the operation of the switch is unidirectional, which prevents any light entering the output ports from affecting the operation of the switch.

Therefore, in order to implement more complex, multi-level switching fabrics, or to handle data-dependent (i.e. self-routing) switching operations, switching elements must have the additional properties and fulfill the more demanding requirements of a logic gate. It is in these applications that most all-optical switching devices are inadequate, as discussed in section 1.4.

1.3.2 Logic Devices

A digital logic gate performs a Boolean operation on one or more binary inputs. All inputs and outputs are of the same physical format thus allowing control to be distributed throughout the switching fabric [3] such that one data stream can control another. The logic gate completely restores signal integrity and timing by replacing the electrons or photons at the input with new electrons or photons from the power supply that go to the output.

Logic levels (i.e. 0 or 1 in digital logic) are physically represented by signal levels, which may be voltage or current for electronics, or for the representationally richer case of optical logic, amplitude, phase, polarization, or color, and are differentiated by a threshold nonlinearity. Information is carried by the logic level and is determined by the interpretation of the signal level (based on the threshold). Since the signal level can be altered during a logic operation, small noise or loss can accumulate throughout the computation and cause information loss. A digital representation of information avoids this degradation if the signal level is restandardized at each step [23]. Restandardization means that the signal levels are restored to values (that are the same throughout the system) which represent valid logic states. As a result, a small deviation about a valid logic level propagates at most one stage and the logic level remains intact (within the allowable noise margin). This is termed logic level restoration.

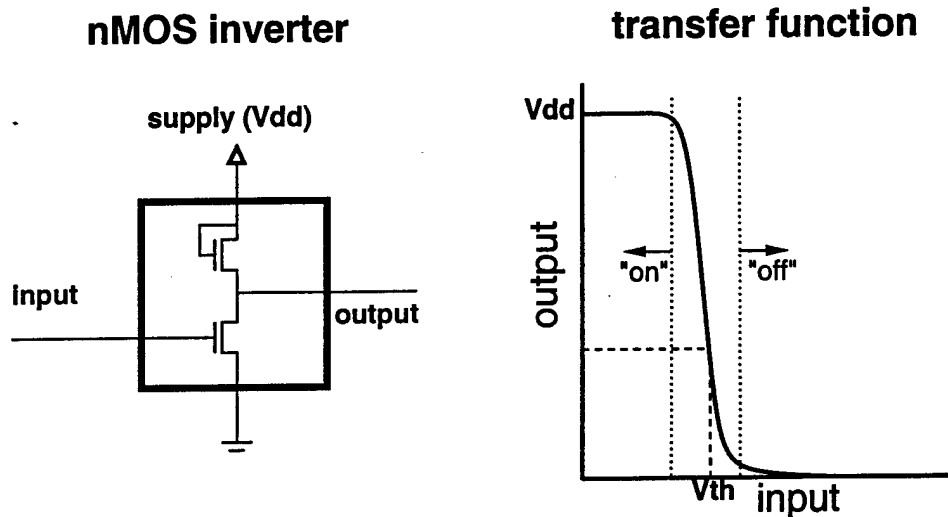


Figure 1.1: nMOS inverter circuit and associated transfer function. Near the threshold voltage is a region of small-signal gain. Large-signal gain is also provided ($V_{dd} > V_{th}$) and the high and low output states are saturated.

The properties required by a digital electronic [24] or optical [25] logic device are well known. Perhaps the most important requirements, upon which many other requirements depend, are gain, saturation, and threshold. These characteristics are illustrated in Figure 1.1, which shows the sigmoidal input-output relationship for a simple nMOS inverter. At the most fundamental level, a region (typically near the input threshold) of differential or small signal gain in the input-output transfer function, where a small change in the input produces a large change in the output, is required (but not sufficient) for a restoring logic gate [26]. Small-signal gain provides sharp switching characteristics and allows for low modulation depth on the input signal biased about the threshold level to produce high contrast switching. Outside the linear region of small-signal gain, saturated levels are key to providing large noise margin by attenuating small variations in the input about valid logic levels (i.e. inputs lying within the "on" and "off" regions shown in the figure), thus producing little change in the output.

Restoration of the logic level prevents accumulation of errors due to attenuation or crosstalk by providing fresh gate output levels directly from the power supply. In the case of an optical gate, logic level restoration must include, in addition to power levels, beam shape and position, pulse duration, color, polarization, and timing [25]. These levels are standardized throughout the entire system typically by the common power supply and ground. Standardization is possible because small-signal gain and saturated levels in the nonlinear transfer function allow for a wide tolerance to the variation in operational characteristics of the devices in the system [24].

Large-signal gain means that the output of the gate is larger than the input required (at least threshold) to switch the gate. Without large-signal gain, fanout, in which the output of a gate provides inputs to many gates in order to implement arbitrary logic functions, is impossible. Most optical logic gates cannot intrinsically provide large-signal gain, and in fact, the output is typically much smaller than the input. Although a separate, external amplification stage can be used, this may limit the bit rate and result in reduced contrast and increased BER when the amplifier noise exceeds the device noise margins. A logic gate with intrinsic gain, in which a small input controls a large power supply, does not suffer this source of noise.

Additional requirements for synthesis of arbitrary logic and switching are logical completeness and cascability. A complete set of logic functions must include inversion [24]. Inversion is a basic function of MOSFET electronics and the inverter structure forms the basis towards implementing more complex Boolean operations such as multi-input and logically complete NOR. Cascability means that the output of one gate can directly drive the input to the next and allows the direct implementation of multi-level logic. Logic level restoration and inputs and outputs of the same format are therefore crucial to allowing cascability.

The most successful logic devices have three orthogonal ports (which can be separate in time, space, wavelength, or polarization) and input-output isolation. A three-terminal device, such as an electronic transistor, ensures that the output of the gate depends only on the inputs and the output does not perturb the operation of the gate [24]. This input-output isolation results in one-way operation and prevents downstream noise which enters the output port from affecting the operation of the gate. There are many examples of two-terminal devices in optics. The problem with two-terminal devices is that they work equally well in either direction [25]. The classic examples are the laser [27] (a two-terminal device with gain) and nonlinear Fabry-Perot etalon [17] which, when critically biased [28], can switch either through the input or through an unwanted reflection back into

the output.

1.3.3 Optical Devices

The main additional requirement for an optical logic device is operation independent of the relative phase between the beam or pulse of light that induces the switching operation and the one that is switched. Phase-sensitivity is not an issue in electronic devices since the phases of the interacting electrons are continuously randomized from phonon collisions [3]. Phase insensitive logic operation is critical in optics because of the coherence of laser sources and since phase is difficult to control in complex systems due to the small optical wavelength. An optical logic gate can be made insensitive to relative phase by using the interaction between optical inputs of orthogonal polarizations or inputs that are spectrally resolved in angle or frequency such that there is no linear interference, but there may be nonlinear interference.

Material nonlinearity is a necessary requirement in order to implement an optical logic gate. The desire for a low loss, ultrafast gate opposes the desire for low switching energy which can be achieved via the use of resonant nonlinearity. Because of the large loss associated with resonant nonlinearity, most ultrafast optical switching and logic gates utilize a much smaller nonresonant nonlinearity. A detailed exploration of the effect of loss on optical soliton logic gates is presented in section 5.3.

It is also desirable for the nonlinear coupling between the interacting optical fields to depend only on intensity, but this can only be guaranteed in certain situations, such as with orthogonal circular polarizations in bulk isotropic media [2, 29]. Other situations can reduce phase-dependent nonlinearity by using the accumulating phase difference between orthogonally polarized linear eigenmodes of naturally or artificially birefringent media to wash out the phase-dependent nonlinear effects.

1.4 Digital Optical Switching and Logic Devices

Some of the most promising all-optical devices for ultrafast switching and logic are the nonlinear directional coupler (NLDC), optical Kerr gate, nonlinear optical loop mirror (NOLM) or Sagnac gate, and the terahertz optical asymmetric demultiplexor (TOAD) and its variants. Even though these devices may possess a region of small-signal gain, only the NLDC has saturated levels in the device input-output curve and none intrinsically provide a robust way to obtain large-signal gain. As a result, the main requirements for these devices are those of a single-stage switch, the most important of which is high contrast, which results from complete switching.

In many of these devices, complete switching occurs upon accumulation of a π differential nonlinear phase shift. This phase is between two independent components of the same beam or pulse and is not to be confused with the relative phase between control and data as discussed earlier on phase-insensitive operation. The accumulated nonlinear phase shift can be written generically

$$\Delta\Phi = \delta \frac{2\pi n_2}{\lambda_f} |A|^2 d, \quad (1.1)$$

where n_2 is the nonlinear Kerr coefficient defined by the total index $n = n_0 + n_2 |A|^2$, λ_f is the free-space wavelength, d is the interaction length, A is the electric field amplitude of the light that induces the switching operation, and δ is a constant factor determined by the nonlinear interaction. The phase shift is proportional to the nonlinear index, the peak intensity, and the interaction length, and can be made large by an increase in any one of these parameters. At the output of the device, this phase shift manifests itself as an amplitude change, either as a rotation of the polarization followed by an analyzer or due to coherent beam combining at an output coupler.

Even though silica has a small nonlinear index for example (see Appendix D), low loss fiber can provide very long interaction lengths d with tight transverse confinement [30], thus making fiber based switching devices very popular [3]. Because switching occurs within the fiber, the control and data must be of different wavelengths and/or polarizations in order to avoid the power loss in coupling in the absence of phase locking and in order to discriminate the switched data from the control at the output. As a result, operation independent of the relative phase between data and control is achieved, but group-velocity or birefringence walkoff limits the interaction length (the walkoff must be less than the pulse durations) and therefore the data rate. These devices are now discussed individually in some detail.

1.4.1 Nonlinear Directional Coupler

This nonlinear switching device is based on the directional coupler as shown in Figure 1.2. The linear directional coupler relies on the coherent interaction between two waveguides placed near each other. The coherent interaction is due to evanescent overlap between fields confined within the individual waveguides and results in periodic energy exchange between the guides. The presence of nonlinearity frustrates this coherent exchange by detuning the waveguides and can result in intensity-dependent

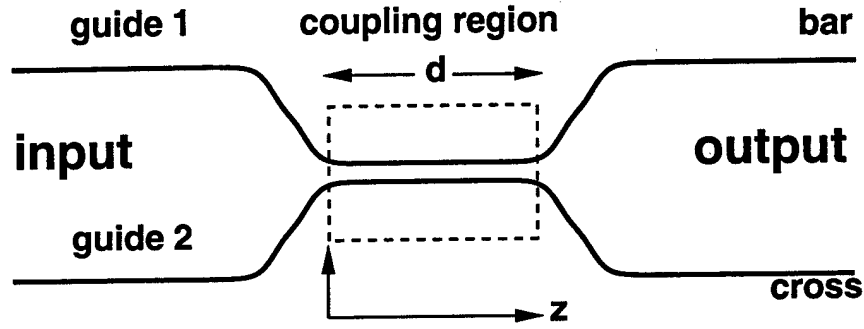


Figure 1.2: Directional coupler. In linear operation, energy is exchanged periodically between guides in the coupling region. Nonlinear operation can frustrate this coupling, resulting in controlled switching.

switching between the output ports. The NLDC was originally proposed [31] as an alternative to bistable switching devices such as the nonlinear Fabry-Perot etalon. More recently, it has received much attention, particularly in the capacity of a prototype all-optical switching device [32], and for ultrafast optical demultiplexing [33].

In the limit of weak coupling between identical guides, the coupled mode equations for the NLDC are

$$-i \frac{\partial A_1}{\partial z} = \kappa A_2 + k_f n_2 |A_1|^2 A_1 \quad (1.2)$$

$$-i \frac{\partial A_2}{\partial z} = \kappa A_1 + k_f n_2 |A_2|^2 A_2, \quad (1.3)$$

where κ is the linear coupling constant, and A_1 and A_2 represent the electric field amplitudes in guides 1 and 2 respectively. Assuming no nonlinearity, the amplitude in each guide can be written

$$\frac{\partial^2 A_1}{\partial z^2} = -\kappa^2 A_1 \Rightarrow A_1(z) = a_1 \cos(\kappa z) + b_1 \sin(\kappa z) \quad (1.4a)$$

$$\frac{\partial^2 A_2}{\partial z^2} = -\kappa^2 A_2 \Rightarrow A_2(z) = a_2 \cos(\kappa z) + b_2 \sin(\kappa z). \quad (1.4b)$$

With the boundary conditions $A_1(0) = A_0$ and $A_2(0) = 0$, the coefficients $b_1 = a_2 = 0$, leading to the result

$$A_1(z) = A_0 \cos(\kappa z) \quad (1.5a)$$

$$A_2(z) = A_0 \sin(\kappa z). \quad (1.5b)$$

The output intensities at the bar and cross ports are given by

$$I_1(d) = I_0 \cos^2(\pi d / L_{\text{coh}}) \quad (1.6a)$$

$$I_2(d) = I_0 \sin^2(\pi d / L_{\text{coh}}). \quad (1.6b)$$

where the coherence length is defined $L_{\text{coh}} = \pi / \kappa$ and is the length over which light is completely coupled from one guide to the other and back. Therefore, for a half-beat length coupler with $d = L_{\text{coh}} / 2$, input into guide 1 emerges at the output from guide 2 (cross state).

In nonlinear operation, using the nonlinear coupled-mode equations, the fractional intensity in each guide is written [31]

$$\frac{I_1(d)}{I_0} = \frac{1}{2} \left\{ 1 + \text{cn} \left(\frac{2\pi d}{L_{\text{coh}}}, \left[\frac{I}{I_{\text{crit}}} \right]^2 \right) \right\} \quad (1.7a)$$

$$\frac{I_2(d)}{I_0} = \frac{1}{2} \left\{ 1 - \text{cn} \left(\frac{2\pi d}{L_{\text{coh}}}, \left[\frac{I}{I_{\text{crit}}} \right]^2 \right) \right\}, \quad (1.7b)$$

where $\text{cn}()$ is a Jacobi elliptic function. As before, this solution assumes that there is no light initially launched into guide 2. The critical intensity is defined as

$$I_{\text{crit}} = \frac{\lambda_f}{n_2 L_{\text{coh}}}, \quad (1.8)$$

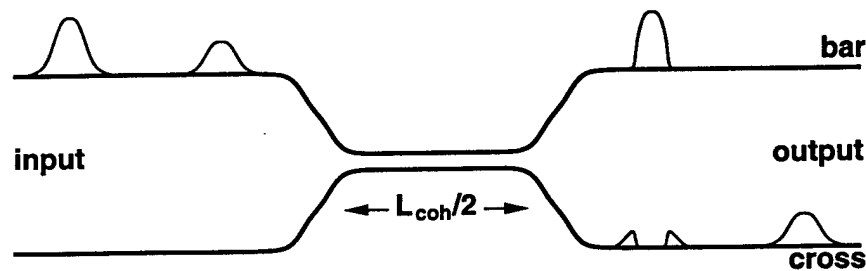


Figure 1.3: Operation of half-beat length nonlinear directional coupler. For low power, or linear, operation, the input pulse exits the cross port (guide 2) while in nonlinear operation, the central portions of the pulse, which exceed the critical intensity, exit the bar port (guide 1) with the wings exiting the cross port.

and represents the input intensity for which the light is divided equally between the two output ports [31]. For a half-beat length coupler, the critical intensity results in a π phase shift for the field in guide 1, and $\delta = 1$ in equation 1.1. Above the critical intensity, the fraction of light exiting the bar port increases towards unity. This two-terminal device performs a switching operation since photons from the data stream are directed into one of two output ports.

A half-beat length NLDC is shown schematically in Figure 1.3, where a low-power data pulse couples to the cross port while a high-power data pulse remains in guide 1 (bar state). The high-power case can be understood intuitively by considering the nonlinear increase in the core index which more strongly guides the light, thus detuning the coupler. Notice that only the central portion of the high-power pulse exceeds the critical intensity and remains in guide 1. This effect is called partial switching, or pulse breakup, and is characteristic of many optical switching devices and results in reduced integrated energy contrast [33]. In fact, partial switching is a problem with all of the non-soliton devices discussed in this section. Solutions to this problem are the use of square pulses [34] such that the intensity is constant across the pulse duration, or the use of temporal solitons [35–37], which propagate unchanged and tend to switch as a unit because of uniform phase across the pulse profile [38].

Figure 1.4 shows a plot of the switching fraction into the bar output port versus normalized peak pulse intensity launched into guide 1 for the half-beat length NLDC. The solid curve shows the fraction of the light that remains in the bar state assuming a constant intensity square-top pulse as given directly by equation 1.7a. The curve for the soliton case has similar shape with sharp switching characteristics [35], but the switching intensity threshold is about twice the critical intensity because the induced nonlinear phase is half that of the plane wave case, as shown in section 2.2. The dashed curve shows the light fraction assuming a non-soliton $\text{sech}^2()$ input intensity profile and takes into account partial switching which leads to reduced contrast operation as indicated by the greatly reduced switching fraction. In either case, at low input levels, all of the light switches over to the cross port, while at high inputs, most of the light exits the bar port.

Direct cascadability can also be addressed using Figure 1.4, where the dotted line represents the input threshold for which the output of the bar port exceeds I_{crit} . Figure 1.5, which shows the transfer function for each port of the device, provides the same information more clearly. As shown in the figure, for a square-top pulse, an input intensity greater than $1.1 I_{\text{crit}}$ is required to exceed the threshold intensity at the bar port. In this case, the output of the bar port can be used as the input to another device resulting in 50% switching of that second device. Therefore, much higher inputs than $1.1 I_{\text{crit}}$ are necessary to cascade to another device, but, due to the lack of large-signal gain and in the absence of external amplification, any optical losses in the system will eventually result in the degradation of the signal level to the point that switching ceases. The situation is worse in the case of a nonuniform pulse. Here, a peak input intensity of greater than $1.6 I_{\text{crit}}$ is required for at least 50% switching of a subsequent device. Indeed, because of partial switching, attempts at cascading two such devices have met with limited success [39]. A final point to note is that when designed as a half-beat-length coupler in linear operation, the NLDC transfer function has quasi-saturated levels which attenuate fluctuation in the input level and allow for some noise margin.

As described so far, the NLDC is a two-terminal device whose switching operation is determined by the intensity of the input signals. A three-terminal version can also be constructed by using a high-power control pulse to switch a low-power data pulse as shown in Figure 1.6, but this device is not directly cascable either due to the lack of large-signal gain and inconsistent signal representation (using different wavelengths). The control pulse must be of different central frequency or polarization in order to avoid phase-sensitivity and to extract the data at the output with zero background. Experimental demonstrations have used different frequencies because of the large difference in coherence lengths between orthogonal polarizations [33]. There are still partial switching problems if the control pulse does not have constant amplitude. An additional problem is dispersive walkoff, which will be discussed in more detail for the Kerr gate. A recent demultiplexing demonstration [33] used a longer control pulse length and allowed the data pulse to walk-through the control to achieve complete switching, and therefore, high contrast. Lengthening of the control pulse means that the separation between data pulses needs to be larger, thereby reducing

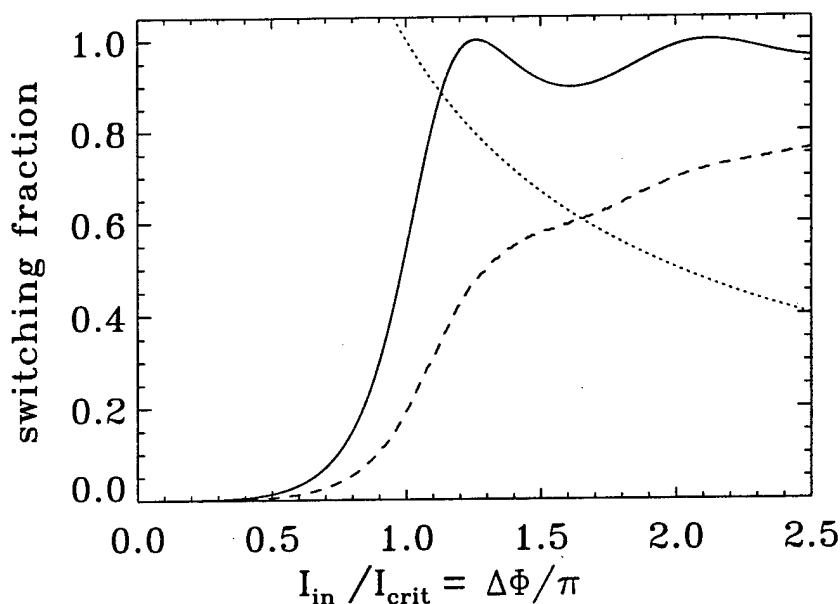


Figure 1.4: Switching fraction into the bar port for a half-beat length NLDC versus input peak intensity. The fraction into the cross port is inverse of the curves shown. The solid curve is for a square-top pulse while the dashed curve is for a non-soliton $\text{sech}^2()$ intensity profile pulse indicating incomplete switching. The dotted curve shows the threshold for which the bar state output exceeds the critical intensity.

the data rate and increasing the switching energy, but also allows for some tolerance to timing jitter.

Because of the region of small-signal gain shown in Figure 1.5, it is possible to obtain large-signal gain with this device, based on the process of light-induced symmetry breaking [40], as originally studied using a phase-dependent interaction between light input into each port. The input into port 1 is biased with a "power supply" or clock pulse near the critical intensity, which is an unstable operating point. Due to small-signal gain, a weak control input into port 2 is amplified, resulting in complete switching into either the bar or cross state depending on the relative phase between the inputs in ports 1 and 2. For a multiple of 2π relative phase, the inputs add constructively and the output emerges from the bar state, while for π relative phase, the inputs add destructively causing the output to emerge from the cross state. This switching is critically dependent on the relative phase between the data and control, and therefore cascability may be difficult to achieve if the phase of the output depends on the intensity of the inputs. It also appears that such a device would be sensitive to variation in waveguide parameters.

A phase-independent version of this three-terminal device with gain can be realized using weak amplitude to initiate switching into the bar port, thereby providing a controlled inversion operation at the cross port. This device must therefore meet the more demanding requirements of a logic gate. Again, if port 1 is biased by a clock pulse at some level less than the critical intensity such that the clock exits the cross port, then additional light of orthogonal polarization or color entering port 1 can activate switching into the bar port. Therefore, logic level restoration is obtained since only the power supply or clock pulse (from the cross port) is passed to subsequent gates when the control pulse is blocked by an analyzer or spectral amplitude filter. For instance, referring to Figure 1.5, for bias in port 1 of $0.8 I_{\text{crit}}$, then an additional contribution of $0.6 I_{\text{crit}}$ (for cross-phase modulation coefficient of $2/3$) can cause complete switching into the bar port. This example illustrates a gain of 1.3, which is the maximum for the standard device and may not be achievable in practice because there is strong variation in switching intensity with power supply and device variations will not allow for level standardization. It is clear that the larger the small-signal gain, the less the input required to switch the state of the device. Because of the ripples in the transfer function, a control pulse of intensity 1.2 will allow for an output level of 0.2 at the cross port, thereby significantly reducing the contrast of the gate. When the switching is incomplete, though, the small-signal gain may be less than unity, precluding the possibility for large-signal gain, as illustrated by the transfer function for the non-soliton $\text{sech}^2()$ (in intensity) pulse.

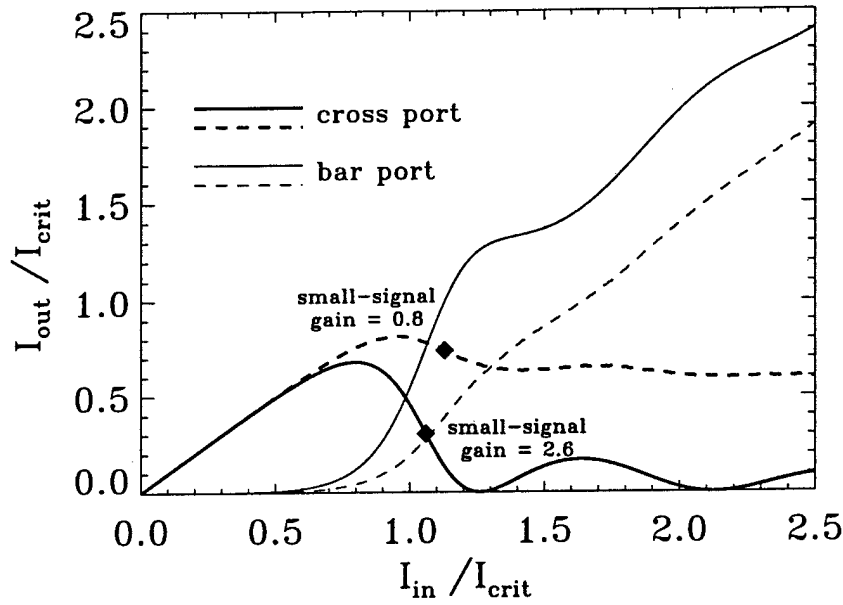


Figure 1.5: Output intensity into the cross and bar states as a function of input peak intensity the half-beat length NLDC. The solid curve is for a square-top pulse while the dashed curve is for a non-soliton $\text{sech}^2()$ pulse.

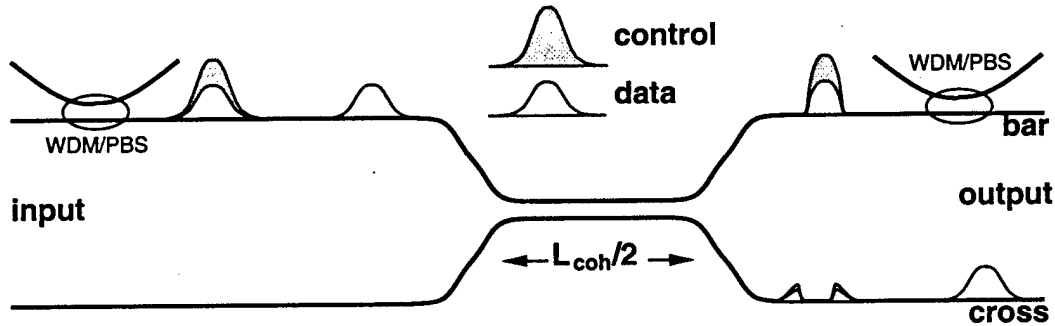


Figure 1.6: Three-terminal half-beat length NLDC. The control pulse is of a different color or polarization than the data pulse.

1.4.2 Kerr Gate

The optical Kerr gate was proposed as an ultrafast photonic gating device [41]. In the implementations discussed here, the Kerr gate is a three-terminal device in which data is either passed or blocked at the output. Since logic-level restoration is not possible, this device is simply an optical switch. The bulk geometry is shown in Figure 1.7. By virtue of the difference between self-phase modulation and cross-phase modulation, the gate beam A_1 induces a differential phase shift which rotates the polarization of the signal beam A_2 , as long as there is nonzero projection of the polarization of A_2 onto a direction orthogonal to the polarization of A_1 . Assuming plane waves interacting at small angles, the total electric field at the input can be written

$$E(0) = \hat{x} [A_1(0)e^{-i\omega_0 t} + \cos\theta A_2(0)e^{-i\omega_0 t} + cc] + \hat{y} [\sin\theta A_2(0)e^{-i\omega_0 t} + cc], \quad (1.9)$$

where θ is the angle that the polarization vector of A_2 makes with the x axis, which is the assumed polarization direction of A_1 . Note that in practice beams must be used in order to spatially separate A_1 and A_2 at the output, where the angle of interaction is greater than the diffraction angle of either beam such that the beams are spatially resolvable in the far field. An additional consideration when dealing with beams (or pulses as in the fiber geometry) is that partial switching occurs analogous to the situation with the NLDC. For simplicity, the plane-wave analysis is used and is sufficient for the present purposes.

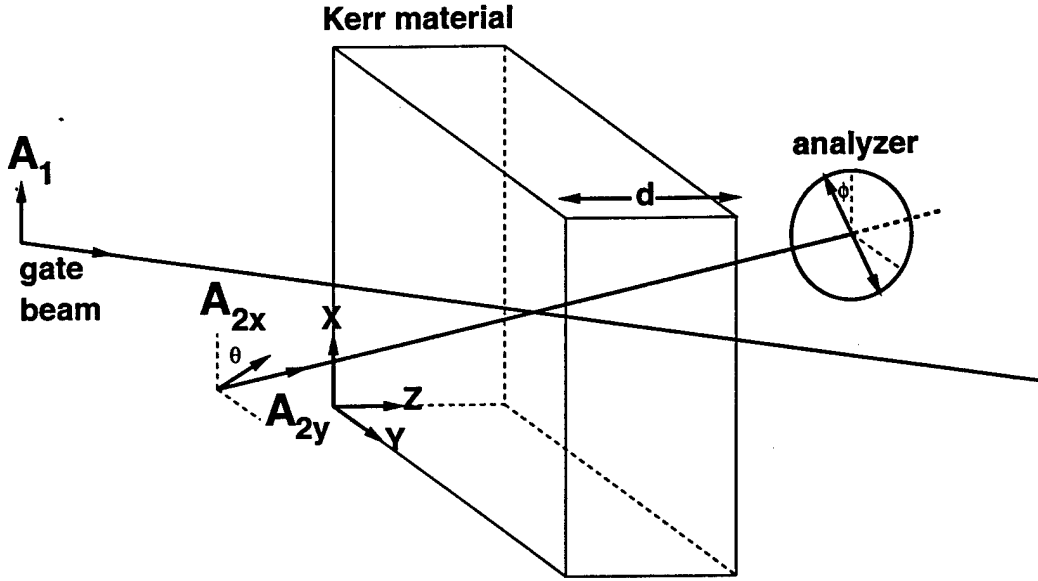


Figure 1.7: Bulk Kerr gate geometry. The polarization of the signal (or data) A_2 rotates based on the intensity of the gate (or control) A_1 . In the absence of A_1 , A_2 is blocked by the analyzer at the output. For A_1 of sufficient intensity, A_2 is transmitted by the analyzer.

After a propagation distance d in the nonlinear medium, the polarization components accumulate linear and nonlinear phase

$$A_1(d) = A_1(0)e^{ik_0d}e^{i\Phi_1} \quad (1.10)$$

$$A_{2x}(d) = A_{2x}(0)e^{ik_0d}e^{i\Phi_{2x}} \quad (1.11)$$

$$A_{2y}(d) = A_{2y}(0)e^{ik_0d}e^{i\Phi_{2y}}, \quad (1.12)$$

where the nonlinear phases are written as

$$\Phi_1 = \frac{2\pi n_2}{\lambda_f} \int_0^d \left[|A_1(z) + A_{2x}(z)|^2 + 2\Delta |A_{2y}(z)|^2 \right] dz \quad (1.13)$$

$$\Phi_{2x} = \frac{2\pi n_2}{\lambda_f} \int_0^d \left[|A_1(z) + A_{2x}(z)|^2 + 2\Delta |A_{2y}(z)|^2 \right] dz \quad (1.14)$$

$$\Phi_{2y} = \frac{2\pi n_2}{\lambda_f} \int_0^d \left[|A_{2y}(z)|^2 + 2\Delta |A_1(z) + A_{2x}(z)|^2 \right] dz. \quad (1.15)$$

The factor 2Δ denotes the cross-phase modulation coefficient and is equal to $2/3$ in isotropic media. Assuming that $|A_1| \gg |A_2|$, thus precluding the possibility for large-signal gain, and that there is no absorption or power exchange throughout the interaction, the nonlinear phase induced on A_2 can be simplified to

$$\Phi_{2x} = \frac{2\pi n_2}{\lambda_f} |A_1(0)|^2 d \quad (1.16)$$

$$\Phi_{2y} = 2\Delta \frac{2\pi n_2}{\lambda_f} |A_1(0)|^2 d, \quad (1.17)$$

with the differential change in phase between the polarization components of A_2 written

$$\Delta\Phi = \Phi_{2x} - \Phi_{2y} = \frac{2\pi n_2}{\lambda_f} [1 - 2\Delta] |A_1(0)|^2 d, \quad (1.18)$$

such that $\delta = 1 - 2\Delta$.

In terms of a Jones vector, the output A_2 before the analyzer can be written

$$A_2(d) = A_2(0)e^{ik_0d}e^{i\Phi_{2y}} \begin{bmatrix} \cos\theta e^{i\Delta\Phi} \\ \sin\theta \end{bmatrix}. \quad (1.19)$$

The transmission of a linear polarizer (analyzer) oriented at an angle ϕ relative to the x axis is [42]

$$T(\phi) = \begin{bmatrix} \cos^2 \phi & \sin \phi \cos \phi \\ \sin \phi \cos \phi & \sin^2 \phi \end{bmatrix}. \quad (1.20)$$

The output intensity of the signal beam after passing through the analyzer is therefore

$$I_2(d) = I_2(0) \left[\sin^2 \phi \cos^2 \phi + \cos^2 \theta \cos^4 \phi + \sin^2 \theta \sin^4 \phi + 2 \cos \theta \sin \theta \cos \phi \sin \phi \cos \Delta \Phi \right], \quad (1.21)$$

where the intensity $I_2(0) = \epsilon_0 c n_0 |A_2(0)|^2 / 2$.

Typically, $\phi = \theta \pm \pi/2$ so that complete extinction is obtained with no nonlinear phase shift, thus implementing a pass gate. In this case, the output intensity is

$$I_2(d) = 2I_2(0) \cos^2 \theta \sin^2 \theta [1 - \cos \Delta \Phi]. \quad (1.22)$$

In order to maximize the contrast of the device, $\theta = \pi/4$, leading to the final result

$$I_2(d) = I_2(0) \sin^2 (\Delta \Phi / 2). \quad (1.23)$$

Therefore, a minimum of a π phase shift is again needed for complete switching, but over switching results when $\Delta \Phi > \pi$. A contrast ratio can be defined based on the difference between the high and low output states of a single port

$$\rho \equiv \frac{I_2(d)|_{\text{high}} - I_2(d)|_{\text{low}}}{I_2(d)|_{\text{high}} + I_2(d)|_{\text{low}}}. \quad (1.24)$$

In the case of the pass gate, $\rho = 1$, assuming an analyzer with perfect extinction.

An inverter can also be realized by setting $\phi = \theta$. The output intensity is

$$I_2(d) = I_2(0) \{1 - 2 \cos^2 \theta \sin^2 \theta [1 - \cos \Delta \Phi]\}. \quad (1.25)$$

Again, $\theta = \pi/4$ maximizes the contrast, leading to the result

$$I_2(d) = I_2(0) [1 - \sin^2 (\Delta \Phi / 2)], \quad (1.26)$$

with $\Delta \Phi = \pi$ for complete extinction of the signal. The contrast ratio of the Kerr gate inverter is

$$\rho = \frac{\sin^2 (\Delta \Phi / 2)}{2 - \sin^2 (\Delta \Phi / 2)} \quad 0 \leq \rho \leq 1. \quad (1.27)$$

Now the contrast is limited by the precision to which the intensity of the gate beam can be controlled in order to exactly achieve the condition $\Delta \Phi = \pi$ for complete extinction of the signal.

Figure 1.8 shows the input-output relationships for the pass gate and inverter geometries. The maximum small-signal gain occurs when $\Delta \Phi = (2m + 1) \pi/2$, but, since $I_1 \gg I_2$, small-signal gain is always much less than unity. In addition, there are quasi-saturated levels for $\Delta \Phi = m\pi$, where m is a positive integer. The most important point to note is that over switching can occur. As $\Delta \Phi$ increases towards π , the pass gate switches on and the inverter switches off, but as $\Delta \Phi$ increases past π , the pass gate switches off and the inverter switches on. Therefore, there are no true saturated levels giving rise to stable states insensitive to variations in the gate beam intensity.

Because of the condition $I_1(0) \gg I_2(d)$, large-signal gain is not intrinsically possible with the Kerr gate either. In other words, it takes a high intensity gating beam to switch a small intensity signal beam. External amplification can be used to bring the high output state after the analyzer to the level required for switching a subsequent stage, but because of the over-switching problems inherent in the device, this process may not be stable and could eventually lead to attenuation of the signal level after a long cascade of devices due to unavoidable amplitude fluctuations. This process can be made stable through the use of a saturating amplifier, but the gain recovery time of the amplifier will place a limit on the data rate.

As an example of the switching intensities needed for the Kerr gate, consider a 1 cm thick sample of fused silica, with $n_2^I = 3.3 \times 10^{-16} \text{ cm}^2/\text{W}$ at $\lambda_f = 1.55 \mu\text{m}$ (see Appendix D). In this situation, a gate intensity of $I_1(0) = 700 \text{ GW/cm}^2$ is required. This intensity can be lowered by orders of magnitude by going to a fiber geometry in which the interaction length can be extended to many km's, as shown in Figure 1.9. For a 1 km fiber gate, $I_1(0) = 3.5 \text{ MW/cm}^2$. Another advantage of the fiber

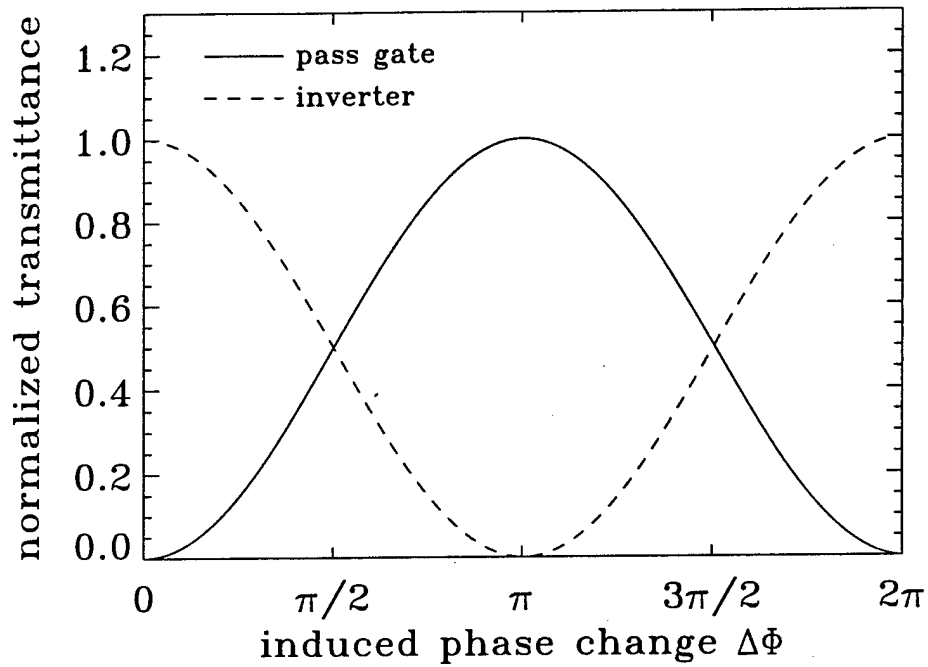


Figure 1.8: Transfer functions for the Kerr gate. The solid curve shows the normalized output versus induced phase shift $\Delta\Phi$ for the pass gate geometry, while the dashed curve shows the same information for the inverter.

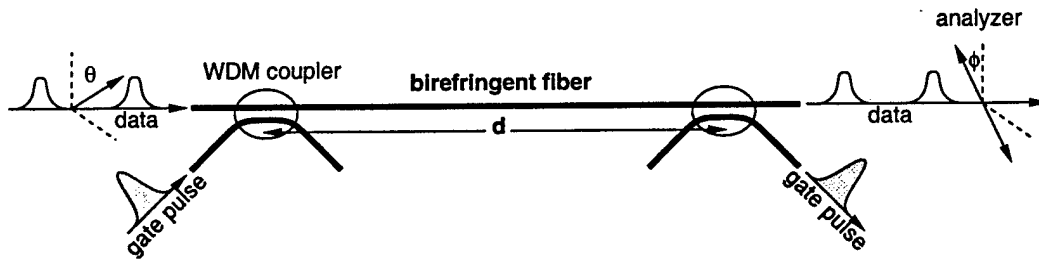


Figure 1.9: Fiber Kerr gate geometry. The gate and data pulses are of different central wavelengths. The gate pulse is timed to co-propagate with and extract one data pulse. Wavelength-division multiplexing (WDM) couplers launch the gate pulse into the interaction region and extract the gate pulse at the output for background-free operation. The bulk analyzer shown can be replaced by an in-fiber polarizer.

geometry in communications situations is that there is no need to couple the signal, or data, pulses out of and back into the fiber in order to achieve the desired switching operation.

Since the pump and data pulses must be separated at the output for background-free operation (i.e. high contrast), the two pulses must be of different color. An additional constraint is that the fiber must be polarization preserving, or birefringent, in order to keep the polarization of the data pulses from randomly evolving due to thermal or stress variations. As a result, the polarization components of the data pulses pick up a nonzero relative phase due to the intrinsic birefringence

$$A_{2x}(d) = A_{2x}(0)e^{ik_x d}e^{i\Phi_{2x}} \quad (1.28)$$

$$A_{2y}(d) = A_{2y}(0)e^{ik_y d}e^{i\Phi_{2y}}, \quad (1.29)$$

where $k_x = 2\pi n_x/\lambda_f$, $k_y = 2\pi n_y/\lambda_f$ and n_x and n_y are the indices of refraction along the principal axes of the fiber. The final polarization state without the gate pulse is in general elliptical, but the original state of polarization can be recovered using a

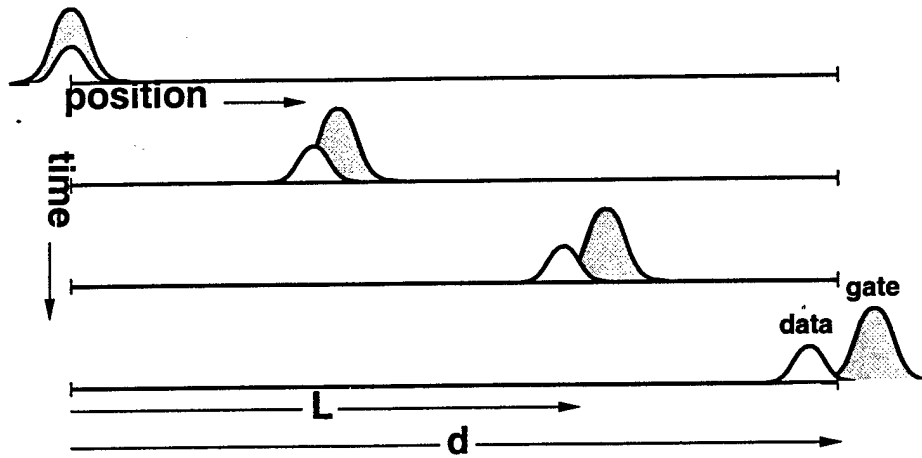


Figure 1.10: Pulse walkoff in fiber Kerr gate geometry. The fiber length is denoted by d , but the interaction length is limited to the birefringent or dispersive walkoff distance L . This diagram illustrates the dispersive walkoff between the gate and same-polarized data. The birefringence walkoff between the two components of the data pulse is similar.

compensator plate. Now, the nonlinear phases are written

$$\Phi_{2x} = 2 \frac{2\pi n_2}{\lambda_f} |A_1(0)|^2 d \quad (1.30)$$

$$\Phi_{2y} = 2\Delta \frac{2\pi n_2}{\lambda_f} |A_1(0)|^2 d, \quad (1.31)$$

where the factor of 2 in the x phase is due to the interaction between waves of different frequency, and the four-wave mixing terms have been neglected. The differential phase is

$$\Delta\Phi = \frac{4\pi n_2}{\lambda_f} [1 - \Delta] |A_1(0)|^2 d, \quad (1.32)$$

which, when $\Delta = 1/3$, is a factor of 2 larger than in the bulk geometry due to the increased cross-phase modulation coefficient between different colors in the same polarization state.

In an early Kerr gate demultiplexing experiment in fiber [43], it was recognized that birefringent and dispersive walkoff would limit the performance of the device due to the long interaction lengths needed to obtain a π phase shift with low switching energies. The result of walkoff is the reduction of the interaction distance as given by the length of the fiber, as shown in Figure 1.10, to an effective interaction distance given by one of the walkoff lengths. If the gate and data pulses have temporal duration τ , then the walkoff lengths can be defined

$$L_{\text{bire}} = \frac{\tau}{|k'_x(\omega_d) - k'_y(\omega_d)|} \approx \frac{c\tau}{|n_x - n_y|} \quad (1.33)$$

$$L_{\text{disp}} = \frac{\tau}{|k'_x(\omega_d) - k'_x(\omega_g)|} \approx \frac{\tau^2}{k''_x(\omega_d)}, \quad (1.34)$$

where $k'_x(\omega_d)$ and $k'_y(\omega_d)$ are the group delay coefficients for the polarization components at the data wavelength, $k'_x(\omega_g)$ is the group delay coefficient at the gate wavelength, $\Delta\omega \sim 1/\tau$ is the approximate spectral separation (by one FWHM), and $k''_x(\omega_d)$ is the group delay dispersion coefficient. For $\tau = 1$ ps pulses (for a data rate ~ 100 Gb/sec with pulse separation 10τ), typical walkoff lengths in silica fiber are $L_{\text{bire}} = 30$ m for $|n_x - n_y| \sim 10^{-5}$ and $L_{\text{disp}} = 35$ m, using $k''_x(\omega_d) = -2.79 \times 10^{-8}$ ps²/μm from Appendix D.

The birefringence walkoff length is the distance over which the two polarization components of the data pulse maintain overlap by at least their temporal FWHM. This is also the distance over which the gate pulse and the orthogonal polarization of the signal pulse maintain overlap. The dispersive walkoff length is the distance over which the gate pulse and the same polarization of the signal pulse maintain overlap. In the worst-case scenario, the gate pulse propagates down the fast (slow) axis of the fiber and has the greater (lesser) group velocity. The interaction distance is therefore limited to the minimum of

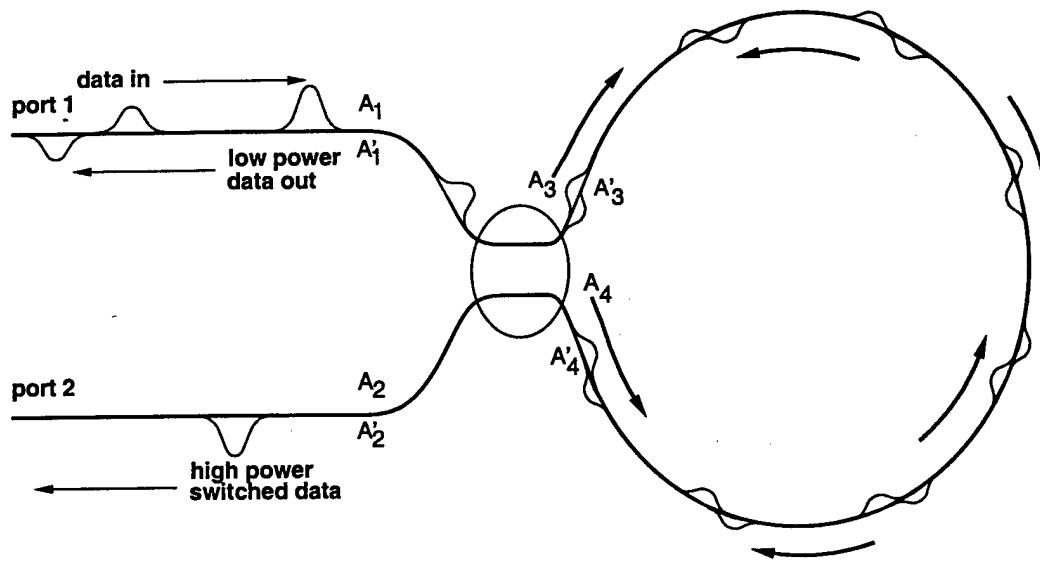


Figure 1.11: Diagram of the two-terminal nonlinear optical loop mirror in the intensity-dependent switching mode.

the birefringent and dispersive walkoff distances. The best-case is for the gate to propagate down the fast (slow) axis and have the lesser (greater) group velocity. Here, the interaction distance is limited to the birefringent walkoff length between the polarization components of the data pulse.

The effective interaction distance can be increased to twice the walkoff length by launching the slower pulse one pulse duration ahead of the faster pulse, but this technique still does not solve the underlying problem. Birefringent walkoff can be effectively eliminated by cross-axis splicing the polarization maintaining fiber such that the distance between each splice is less than the walkoff length L_{bire} [43]. The effects of dispersive walkoff can be reduced by decreasing the spectral separation between the data and gating pulses, thus limiting the data rate because $1/\tau \sim \Delta\omega$, or by equally spacing the central wavelengths about the zero dispersion wavelength such that both pulses travel with the same group velocities. This does not allow for both pulses to be temporal solitons, however.

A recent fiber Kerr gate demonstration used a 1.5 cm length of silica fiber (which is about L_{disp} for the experimental pulse duration) to perform demultiplexing of a 460 Gb/s data stream [44]. In the experiment, the peak gate intensity 43 GW/cm² (about 5 nJ pulse energy) resulted in a gating efficiency of less than 10%. A π phase change requires peak gate pulse intensity of nearly 235 GW/cm² (about 25 nJ pulse energy), which, except for the benefits of compactness and increased coupling efficiency, eliminates the original advantages of going to the fiber geometry.

1.4.3 Nonlinear Optical Loop Mirror

The nonlinear optical loop mirror [38], or NOLM, has been used extensively in ultrafast all-optical demultiplexing experiments [45, 46]. Studies have also focused on the use of NOLMs to realize simple logic gates [45] and more complicated logic circuits [16]. The most basic two-terminal configuration is shown in Figure 1.11. A pulse which enters port 1 is split at a 2x2 fiber coupler into two counter-propagating pulses which recombine at the coupler. For a perfectly symmetric device, the pulses constructively interfere at the coupler and exit from port 1, thus acting as a mirror. Deviations from symmetry result in light exiting both ports 1 and 2. In nonlinear operation of a non-symmetric device, the pulse intensity can control the ratio of light exiting each output port.

Immediately after the coupler, the field amplitude in each counter-propagating direction of the loop can be written

$$A_3 = tA_1 \quad A'_3 = A_4 e^{i\phi_{\text{ccw}}} \quad (1.35)$$

$$A_4 = irA_1 \quad A'_4 = A_3 e^{i\phi_{\text{cw}}} \quad (1.36)$$

when it is assumed that no light enters port 2. The unprimed quantities represent the amplitudes before propagation around the loop while the primed quantities are after propagation around the loop and just before the coupler. The phases around the loop are allowed to be different as justified later. Here t is the (real) amplitude transmission coefficient and r is the (real) amplitude reflection coefficient. When the angle of incidence is less than the critical angle (as it must be here), there is a π phase shift upon

reflection from one of the coupler interfaces. Since the 3dB fiber coupler is symmetric, this phase is distributed symmetrically with the reflection coefficient in each direction. The amplitudes at the output ports are then

$$A'_1 = tA'_3 + irA'_4 = itrA_1 e^{i\phi_{ccw}} [1 + e^{i\Delta\Phi}] \quad (1.37)$$

$$A'_2 = tA'_4 + irA'_3 = A_1 e^{i\phi_{ccw}} [t^2 e^{i\Delta\Phi} - r^2], \quad (1.38)$$

where $\Delta\Phi = \phi_{cw} - \phi_{ccw}$. Finally, the output intensities are written

$$I'_1 = 2r^2[1 - r^2][1 + \cos \Delta\Phi] I_1 = 4I_1 r^2 [1 - r^2] \cos^2(\Delta\Phi/2) \quad (1.39)$$

$$I'_2 = I_1 - 4I_1 r^2 [1 - r^2] \cos^2(\Delta\Phi/2), \quad (1.40)$$

assuming a lossless system such that $r^2 + t^2 = 1$.

For the original NOLM device which switches based on the intensity of the incoming data pulses, the phases are

$$\phi_{ccw} = k_0 L + \frac{2\pi n_2}{\lambda_f} \int_0^L |A_4(z)|^2 dz \approx k_0 L + \frac{2\pi n_2 L}{\lambda_f} |A_4(0)|^2 \quad (1.41)$$

$$\phi_{cw} = k_0 L + \frac{2\pi n_2}{\lambda_f} \int_0^L |A_3(z)|^2 dz \approx k_0 L + \frac{2\pi n_2 L}{\lambda_f} |A_3(0)|^2, \quad (1.42)$$

where the differential phase is written

$$\Delta\Phi = \frac{2\pi n_2 L}{\lambda_f} [|A_3(0)|^2 - |A_4(0)|^2] = \frac{2\pi n_2 L}{\lambda_f} [1 - 2r^2] |A_1(0)|^2, \quad (1.43)$$

and $\delta = 1 - 2r^2$. For a symmetric device in which $r^2 = 0.5$, the phase shift is zero and all the light exits port 1, independent from the strength of the nonlinear interaction. Therefore, the loop must be unbalanced in order for self-switching to occur, in which the light exits port 2.

An example of the self-switching of the NOLM is shown in Figure 1.12, which plots the peak intensity in each of the two output ports (assuming square-top pulses) using $r^2 = 0.25$. This figure is different than Figure 1.8 of the Kerr gate, which plots the switching fraction, not the output intensity. This is done here to emphasize the low contrast of the two-terminal NOLM. Like the Kerr gate, over switching can occur for input intensity beyond that required for $\Delta\Phi = \pi$, thus making this device difficult to cascade. It should also be noted that partial switching will occur if square-top or soliton pulses are not used.

Because the self-switching behavior is similar to the two-terminal NLDC, a three-terminal NOLM with large-signal gain can be constructed in the same way, by biasing the device with a strong power supply or clock pulse and using a weaker control or data pulse (of different polarization or color) to initiate switching, thereby providing complete logic level restoration. Unlike the NLDC, the NOLM switching function does not possess saturated levels, thus making the device sensitive to input fluctuations, but also allowing for multiple intensity levels for the clock pulse (spaced by even intervals on the figure) to achieve inverting operation. Because the switching intensity is fixed at $1/2\Delta$ for the coordinate system used in the figure, the large-signal gain can be arbitrarily high (subject to limitations placed by nonlinear absorption, for example). The output from either port 1 or port 2 can be used. The contrast from port 1 is unity, but since this is also the input port, a coupling loss must be taken both at the input and at the output in order to extract the switched data.

The contrast of the NOLM at output port 2 is

$$\rho = \frac{2r^2[1 - r^2]}{1 - 2r^2[1 - r^2]}, \quad (1.44)$$

which is maximized when $r^2 = 0.5$. The peak intensity required for a π phase shift is

$$I_1 = \frac{\lambda_f}{2n_2 L [1 - 2r^2]}, \quad (1.45)$$

which is infinite when $r^2 = 0.5$. Here, the intensity I_1 represents the data pulse intensity plus any additional control pulse intensity. For example, $I_1 = I_{\text{data}} + 2\Delta I_{\text{control}}$, in the case of the biased device. The minimum required intensity is when $r = 0$, but the contrast is 0. Therefore, in the two-terminal or biased three-terminal NOLM, there is a tradeoff between high contrast operation and low switching intensity (energy).

Some of these problems can be solved by using a balanced ($r^2=0.5$) and un-biased three-terminal NOLM [47]. This configuration is shown in Figure 1.13, in which an external control pulse of different polarization or wavelength is used to

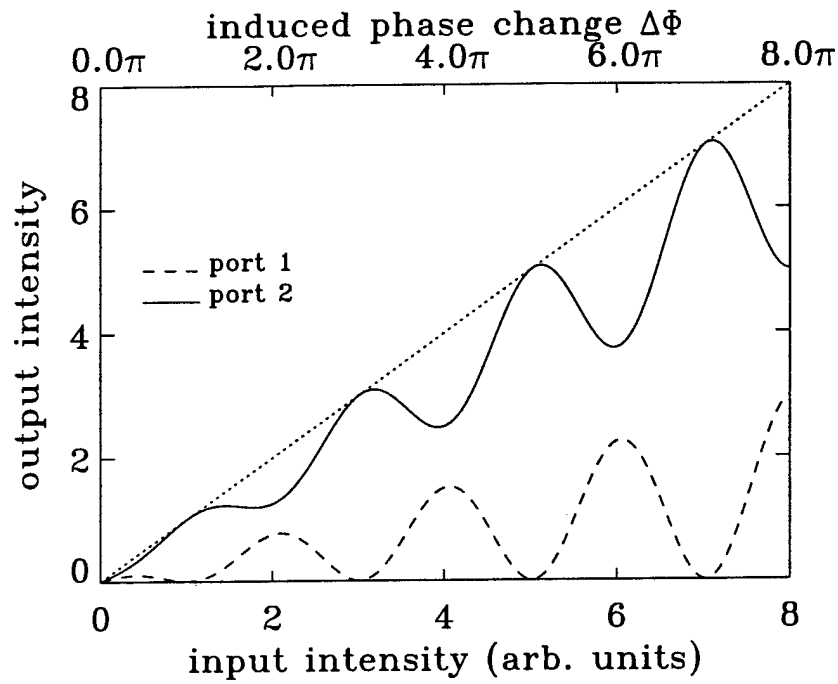


Figure 1.12: Switching characteristics of the two-terminal NOLM. If a non-square-top or non-soliton pulse is used, the output energy will not follow the intensity curve and the energy contrast will be reduced over the ideal case of unity for port 1 or as given by equation 1.44 for port 2. The amplitude reflection coefficient of the 2x2 fiber coupler is set to $r = 0.5$, giving a reflectance of $r^2 = 0.25$.

switch a much weaker data pulse by inducing a π phase shift on the counter-clockwise pulse via cross-phase modulation. In this case, the phases of the counter-propagating pulses can be written

$$\begin{aligned}\phi_{ccw} &= k_0 L + 2\Delta \frac{2\pi n_2 L}{\lambda_f} |A_{ctl}(0)|^2 \\ \phi_{cw} &= k_0 L,\end{aligned}\tag{1.46}$$

where A_{ctl} is the amplitude of the control pulse, which co-propagates A_{ccw} and therefore counter-propagates A_{cw} , inducing negligible nonlinear effect on the latter. For orthogonal polarizations, $2\Delta = 2/3$, while for different wavelengths, $2\Delta = 2$. In this device, the switching intensity and contrast are independent of r , because the change in phase is

$$\Delta\Phi = 2\Delta \frac{2\pi n_2 L}{\lambda_f} |A_{ctl}(0)|^2\tag{1.47}$$

and the contrast at output port 2 is $\rho = 1$ assuming that the control pulse intensity can be maintained to achieve $\Delta\Phi = \pi$.

Due to the long interactions lengths of 10 km or more [46], these switching elements suffer from polarization or dispersive pulse walkoff (unlike the fiber Kerr gate, the control pulse does not have to be of different polarization *and* wavelength). Walkoff can be minimized by using cross-spliced polarization maintaining fiber or choosing the control and data wavelengths symmetrically about the zero dispersion wavelength [45], as discussed for the fiber Kerr gate. Recent experiments have achieved demultiplexing of 100 Gbit/s data streams [48] in a 6 km loop using 1 pJ control pulses, and two-stage cascading using low birefringence fiber loops with an amplification stage in between [49].

1.4.4 Terahertz Optical Asymmetric Demultiplexor

The terahertz optical asymmetric demultiplexor [50], or TOAD, was developed to reduce the long latency inherent in the NOLM and reduce switching energies by asymmetrically positioning a thin, highly (resonant) nonlinear element on one side of center of a fiber loop mirror, as shown in Figure 1.14. The fiber nonlinearity is not utilized and therefore the fiber loop can be on the order of meters in length rather than kilometers.

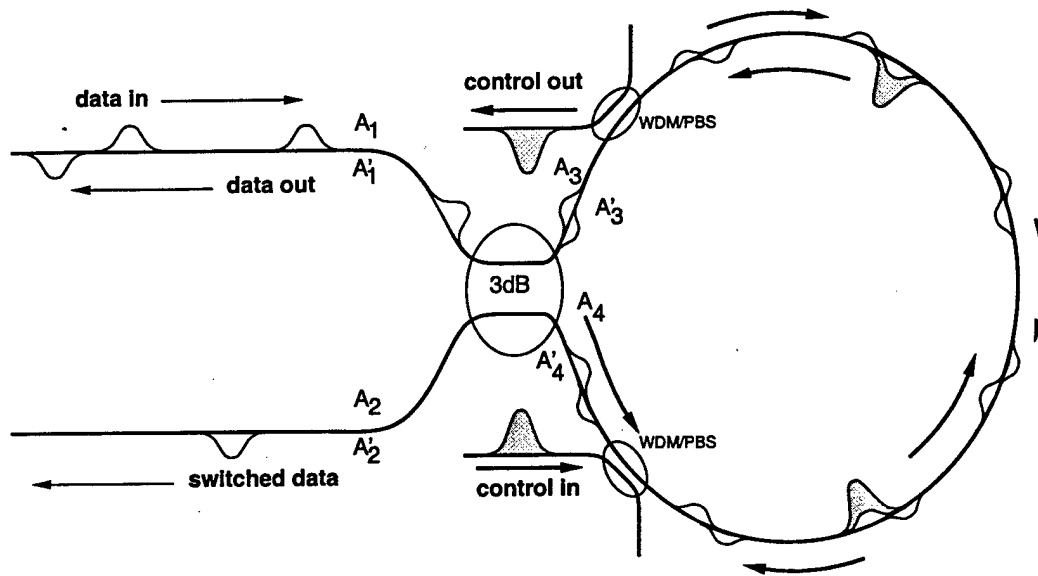


Figure 1.13: Three-terminal NOLM in two-pulse configuration. Wavelength division multiplexing (WDM) or polarizing (PBS) couplers transfer the control pulse into and out of the loop with nearly unity efficiency. For polarizing couplers, the fiber loop must be polarization maintaining. The control pulse only propagates counter-clockwise, and only interacts strongly with the counter-clockwise component of the data pulse.

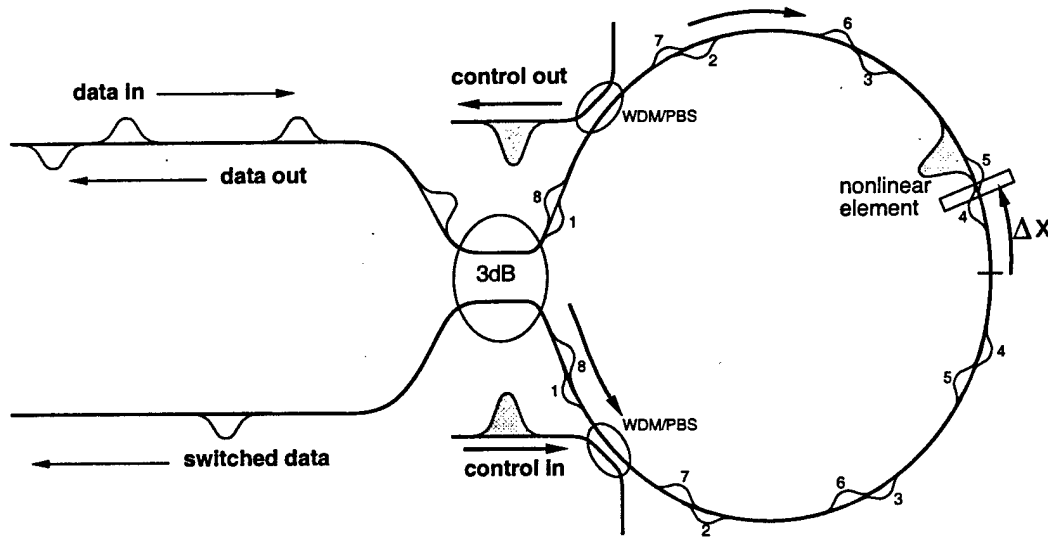


Figure 1.14: Three-terminal terahertz optical asymmetric demultiplexor. The nonlinear phase shift is provided by the control pulse excitation of an asymmetrically placed nonlinear element.

The TOAD switch works similarly to the three-terminal NOLM by imposing a π phase shift in one arm (direction) of the loop, and cannot be configured as a true logic gate. The differentiating factor from the NOLM is that a thin nonlinear element excited by the control pulse is responsible for the nonlinear phase shift induced on the data pulses. In the diagram, pulses 1-3 pass through the nonlinear element before the control pulse and experience no nonlinear phase shift in the clockwise and counter-clockwise parts because the nonlinear element is in the un-excited state. As a result, there is no differential phase shift ($\Delta\Phi = 0$) and the pulses exit the input port (or reflect). Pulses 5-8 pass through the element after the control pulse, but now each direction experiences a nonlinear phase shift, but, given constraints of long relaxation time of the nonlinear element as discussed later, experience no significant differential phase shift and also exit the input port. Pulse 4, on the other hand, switches out the other port because the clockwise propagating part of the pulse passes through the nonlinear element before the control pulse and experiences no nonlinear phase shift, but the counter-clockwise part reaches the nonlinear element just after

excitation by the control pulse and experiences a nonlinear phase shift. If the differential phase shift $\Delta\Phi = \pi$, then the pulse will exit the output port.

The counter-clockwise and clockwise phases can be written

$$\phi_{ccw} = k_0 L + \Delta \frac{2\pi d}{\lambda_f} \Delta n(t_i - t_{ctl} + 2\Delta x/v_g) \quad (1.48)$$

$$\phi_{cw} = k_0 L + \Delta \frac{2\pi d}{\lambda_f} \Delta n(t_i - t_{ctl}), \quad (1.49)$$

where t_i is the time that the clockwise portion of the i^{th} pulses reaches the nonlinear element, $t_i + 2\Delta x/v_g$ is the time that the counter-clockwise portion of the same pulse reaches the nonlinear element, t_{ctl} is the time that the control pulse excites the nonlinearity, and v_g is the group velocity. The time-dependent nonlinear index change is written

$$\Delta n(t) = \int_0^\infty R(\tau) |A_{ctl}(t - \tau)|^2 d\tau \quad t \geq 0 \quad (1.50)$$

$$= 0 \quad t < 0, \quad (1.51)$$

where $R(\tau)$ is a nonlinear response function, typically of the form of an exponential decay. It is clear that, assuming an instantaneous rise time,

$$t_i < t_{ctl} < t_i + 2\Delta x/v_g, \quad (1.52)$$

in order to switch the i^{th} pulse. A much more complete analysis has been performed [51], with the basic conclusion that the nonzero length of the nonlinear element places a restriction on the minimum duration of the switching window, which must be at least twice the transit time through the element plus the rise time of the nonlinearity. The relaxation time of the nonlinearity sets the lower bound on the frame time because the nonlinear material must be in the ground state before re-excitation by another control pulse. These requirements are expressed by [50]

$$\tau_{rise} + 2\tau_{transit} < T_{bit} \ll \tau_{fall} < T_{frame}, \quad (1.53)$$

where τ_{rise} is the rise-time of the nonlinearity and can usually be considered instantaneous, $\tau_{transit}$ is the transit time through the nonlinear element, T_{bit} is the time between adjacent bits in the data stream, τ_{fall} is the fall time of the nonlinearity and does not have to be instantaneous as in the other devices which require $\tau_{fall} \ll T_{bit}$, and T_{frame} is the frame time for time-division multiplexed data. The fall time τ_{fall} must be greater than the time difference $2\Delta x/v_g$ between when the two parts of the same pulse pass through the nonlinear element so that the clockwise and counter-clockwise portions accumulate approximately the same nonlinear phase shift. A disadvantage is that only one data pulse can be switched within the relaxation time of the nonlinear element, so that the switching events must occur at least τ_{fall} apart. Unlike the other devices, the TOAD is relatively insensitive to timing jitter between the control and data pulses. All that is required is that the control pulse arrives (up to one bit time) at the nonlinear element before the intended data pulse. The amount of timing jitter that can be tolerated is limited to $T_{bit} - \tau_{rise}$.

Because the nonlinearity need not be instantaneous (and in fact the fall time must be long), very large, resonant nonlinearity, such as that produced by real particle excitation and decay, can be used. The optical path length of the nonlinear element must be less than that between the element and center and, since the nonlinear phase shift is proportional to the product of nonlinear index, control pulse intensity, and the thickness of the nonlinear element, the latter is perhaps the main limitation on high contrast, low energy switching for THz data streams. A recent demonstration [52] showed single-pulse demultiplexing of 250 Gbit/s data with 0.8 pJ switching energy.

A related device, based on a Mach-Zehnder configuration with asymmetric placement of a nonlinear element in each of the two arms (where the difference between the distance from the coupler to each element is Δx), has also been demonstrated [53]. As in the TOAD device, a control pulse is used to excite the nonlinear elements, where the time difference between the excitation of each element leads to a switching window of duration $\Delta x/v_g$, but again, large-signal gain cannot be realized. Because the two arms are physically distinct, any differential environmental variations can disturb the operation of the gate, but the effect is less important for an integrated device.

Another variation, which perhaps is more akin to the fiber Kerr gate, is the so-called ultrafast nonlinear interferometer [54], or UNI. This device is shown in Figure 1.15. The input data pulses split equally along the polarization axes of birefringent fiber. Because the group velocity is different along each axis, the polarization components walk-off from each other in time. The time difference between the leading and trailing components allows for asymmetric excitation of the nonlinear element by a strong gate pulse such that the leading pulse is affected only by the amplification process of the device (the device is electrically pumped) and the trailing pulse is affected by the device after gain saturation by the strong gate. Therefore, due to sub-ps carrier

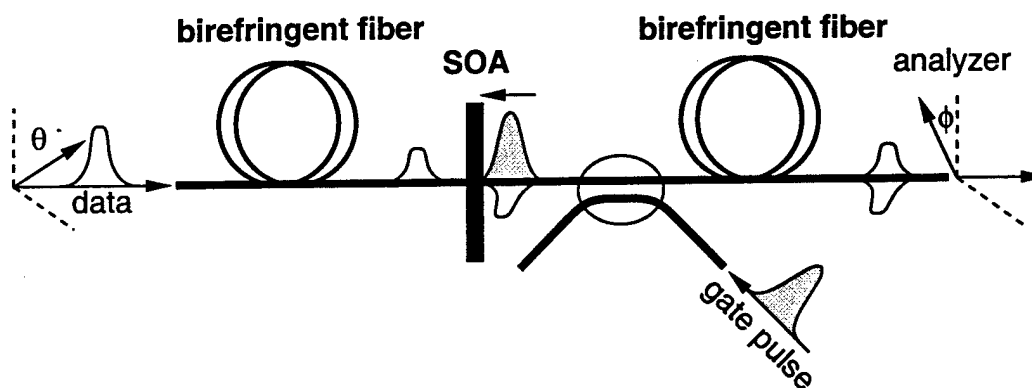


Figure 1.15: Three-terminal ultrafast nonlinear interferometer. A differential phase shift is imparted onto the data pulse at the semiconductor optical amplifier (SOA) by asymmetric excitation with the gate pulse.

heating and gain depletion, and virtual electronic processes, both differential amplitude and phase changes are induced between the data components. Long-lived carrier population and thermal nonlinearities ensure that no differential change is induced between subsequent data pulse components. After traveling through another length of birefringent fiber with fast and slow axes reversed, the components overlap at an analyzer biased to produce the desired logic operation. For example, if the analyzer is set to pass the data pulse in the absence of the gate pulse, then the device will act as an inverter, albeit without logic level restoration and large-signal gain.

Like the TOAD, the switching rate is limited by the gain recovery time of the SOA, but it should be noted that only enough recovery time is needed so that the control pulse can induce another π phase shift. If only TDM de-multiplexing is performed, then recovery time only limits the frame period, while the bit rate is limited by the optical thickness of the element and the rise time of the nonlinearity. For switching operations that occur on a bit-by-bit basis, then the π phase shift recovery time places the main limitation on bit rate, which can be as high as 100 Gb/s [55]. A recent experiment [11] demonstrated the cascability of this three-terminal device by implementing a 40 Gb/s inverting shift register. A fiber amplifier was used to bring the inverted data output of the device up to the necessary 0.1 pJ switching energy.

1.5 Optical Soliton Logic Devices

The devices discussed in the previous section all suffer from drawbacks that may limit their use in cascaded digital logic and switching systems. Except for variations of the nonlinear directional coupler (NLDC) and nonlinear optical loop mirror (NOLM), none of these switching devices has complete logic level restoration with large-signal gain. In the standard configurations, the data pulse itself is transmitted to the output. The strong control pulse only initiates the switching and is thrown away afterwards. As a result, the transmitted data is never restored allowing signal level variations to build up over time. The roles of control and data can be reversed such that weak "power supply" pulses are passed on to subsequent gates, and are controlled by strong data pulses. Here, restoration is almost complete, with the lack of obtaining a valid signal level without external amplification which introduces ASE noise and bandwidth limitations. The biased three-terminal NLDC and NOLM have complete logic level restoration because of large-signal gain, but the gain of the NLDC is very limited and suffers from low contrast in practice, and the NOLM does not possess saturated levels and is lossy in order to achieve high contrast.

In the three-terminal configurations, these devices are only directly cascable when using control or gate pulses of orthogonal polarization to the data pulses. Cascadability is achieved simply by using a half-wave plate to rotate the polarization of the output to serve as a control input. Using different colors also allows for cascability, but the system becomes more complicated. In the first realization, the output of the gate must be converted through some nonlinear process such as three- or four-wave mixing, to the appropriate control input color. A second realization is to alternate the correct control color at each stage such that no wavelength conversion is necessary, but this may require two device designs in order to achieve optimum performance. These two-color schemes are the only way to cascade the fiber Kerr gate, since different colors are necessary in the fiber geometry, and the NLDC, because the use of different polarizations greatly reduces the switching performance [33].

The Kerr gate, NOLM, and terahertz optical asymmetric demultiplexor (TOAD) all suffer from over switching, although saturated gain in the nonlinear element can clamp the total induced phase shift in the latter device. In each of these three devices, a control intensity higher than that required for a π phase shift will cause over switching [45] and result in reduced contrast. Therefore, each gate in the system must be precisely toleranced if high contrast operation is required, but because of

unavoidable variations, the signal level will be attenuated through a long cascade with eventual loss of the logic level.

The performance of all of the devices, except TOAD, is reduced considerably by walkoff and incomplete switching. Means of reducing walkoff were discussed for each device, but it should be noted that, since the use of orthogonal polarizations is more practical for cascability considerations, cross-axis splicing is an elegant solution. Because square-top pulses possess high frequency components at the leading and trailing edges leading to rapid pulse broadening, the most robust way to eliminate incomplete switching is to use temporal solitons for the control and data pulses. The use of temporal solitons can also reduce or eliminate walkoff by the mechanism of temporal trapping [56,57].

The limitations of these devices motivates the study of a new class of devices based on optical solitons, which make full use of the soliton properties. Because solitons are threshold phenomena, they are natural carriers of digital information. Below a critical power or energy, a nonlinear wave will asymptotically disperse, but above the critical power or energy, a nonlinear wave will evolve into one or more stable (possibly higher-order) solitons. Therefore, binary logic levels can be represented by the amount of detected power or energy, or even by spatial or temporal size (a soliton could readily pass through a spatial or temporal aperture, while the majority of a dispersive wave would not). A related issue is resolvability. Since solitons beat the linear diffraction/dispersion limits, a logic gate (without absorption) could be arbitrarily long. This property is used to great advantage in the soliton collision and dragging gates as discussed later, where a small angular/frequency change (induced by a nonlinear phase shift less than π) manifests itself as a resolvable spatial/temporal shift.

Other benefits arise from the use of solitons in digital logic systems. Solitons are stable to weak perturbations meaning that the soliton tends to maintain its shape in the presence of material inhomogeneities or input profile variations, which is an important property for logic restoration. In fact, the original amplitude and shape can be restored even after absorption by the use of adiabatic amplification. This property is useful for logic restoration and in long haul communications. Additionally, in some cases, solitons will maintain their shape even after collision with another soliton, which turns out to be a disadvantage for optical switching. The logic gates presented in this thesis use a configuration in which the interaction is inelastic such that soliton shape is not necessarily maintained after collision. The most important benefit to optical switching is that solitons (or nonlinear waves in general) can exert a force on one another, unlike light in linear propagation. These forces can alter the direction in space or velocity in time of one or both solitons resulting in a switching or logical operation.

In order to achieve low switching energy per gate operation, the ideal optical soliton logic gate should be based on one of the three geometries that allow for complete three-dimensional confinement [2]: 1-D temporal solitons in fiber, 2-D spatio-temporal solitary waves in slab waveguides, and 3-D light bullets. The advantage of spatio-temporal solitary waves over temporal solitons for switching and logic is the removal of one or two transverse dimensions of linear confinement. This allows for spatial parallelism (without the loss of ultrafast temporal pipelining) and the freedom to use spatial interactions with much easier output state discrimination [2,58] than their temporal counterparts [59] in order to implement amplitude keyed logic for eventual conversion to the electronic domain. Even though one dimension of spatial parallelism is lost, the 2-D solitary wave has the important advantages over the 3-D "light bullet" case in that it should be more easily realizable experimentally [60], and the slab waveguide geometry allows for the possibility for tailoring of the dispersion properties [61], photolithographic circuit definition, and cooling through the area of the substrate. The (2+1)-D case also presents an efficient test bed allowing for the study of the effects of higher-order terms in the evolution equation by retaining full spatio-temporal dynamics without resorting to the time-consuming (3+1)-D simulation.

The remainder of this section briefly discusses two soliton-based logic gates based on the trapping interaction. These gates satisfy the requirements for a logic device, including large signal gain and logic level restoration. The first device is the temporal soliton dragging gate [62], and the second is the related spatial dragging gate [58], which is the main subject of this thesis. Note that other soliton interactions can be used for logic devices and are discussed further in Chapter 5.

1.5.1 Temporal Soliton Dragging Gate

The soliton trapping [62] and dragging [59] gates are based on the temporal trapping mechanism [56,57]. Two nonlinear waves of different polarization and/or wavelength propagate down a fiber with different group velocities. If they are initially overlapping in time, in linear propagation, one will reach the end of the fiber before the other. In nonlinear propagation though, they can trap each other through nonlinear attractive forces such that they both propagate at the same group velocity. In this case, each wave must experience a frequency shift in order to propagate at the common, weighted-mean, group velocity. The two nonlinear waves do not need to be of the same size (in terms of pulse area) in order for trapping to occur [57], and the frequency shift experienced by the smaller wave is greater than that experienced by the larger wave.

Figure 1.16 shows the generic three-terminal gate geometry. In the most basic geometry as shown, these gates perform an inversion operation in which the pump is passed in the absence of the signal and blocked by a spectral filter or time gate (as discussed later) in the presence of the signal. Note that the signal is always blocked at the output and the pump of one gate becomes the signal of the next, resulting in true three-terminal operation. Since the pump pulse may be passed on to later gates,

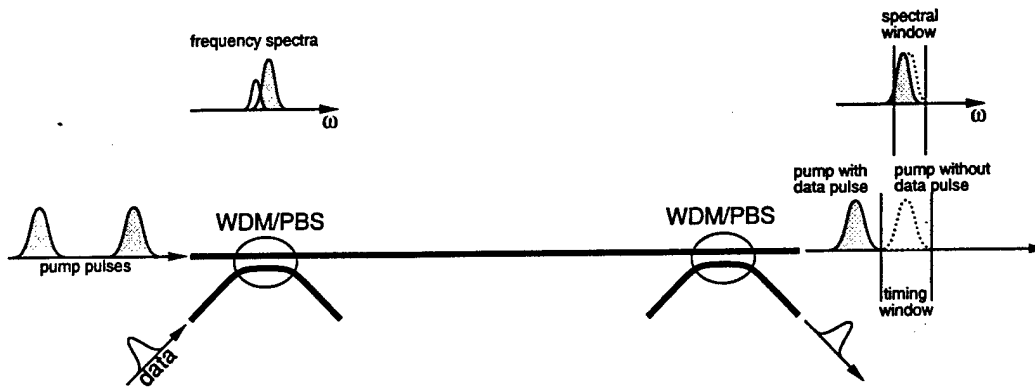


Figure 1.16: Temporal soliton trapping/dragging gate. Soliton pump pulses (always present) initially overlap in time with (non-soliton) data or signal pulses of different wavelength and/or polarization. Cross-induced chirp causes the pump and signal to propagate with the same group velocity. This frequency shift must be resolvable for the trapping gate. The change in group velocity becomes an arbitrarily large timing shift after propagation through a dispersive fiber delay line. For the dragging gate, this timing shift must be resolvable.

it is convenient for it to be a temporal soliton in order to maintain its shape after propagation through many tens to hundreds of meters of fiber, thereby ensuring logic level restoration. The trapping and dragging gates use pulses of the same wavelength but with orthogonal polarizations. Note that similar type gates can be constructed using pulses of different wavelength [63], based on the analogous trapping mechanism [64], but this gate is not readily cascable because frequency shifters are necessary to drive the subsequent gate, while in the case of different polarizations, a wave plate can be used to rotate the output pump polarization to the correct state. In either case, the gate operation is independent of the relative phase between the pump and signal¹.

The soliton trapping gate relies on a resolvable spectral shift of the pump soliton [62], so that an unshifted pump will pass through a spectral bandpass filter while a shifted pump (in the presence of the signal) will not. This results in amplitude keyed logic and is compatible with common high-speed optical detectors for eventual conversion into an electronic signal. In order for this to occur, the fiber birefringence must be sufficiently large to produce the necessary difference in group velocity between the pulse propagating down the slow axis and the pulse propagating down the fast axis, which when compensated due to trapping, results in a spectrally resolved shift. The pump and signal pulses must also be of nearly the same amplitude so that the shift is not weighted preferentially towards the signal. As a result, this gate cannot provide significant large-signal gain. In fact, the first experimental demonstrations [62, 65] used solitons of 300 fs duration and 42 pJ energies for the pump and signal. A final note is that, because complete trapping can occur in a few soliton periods, the gate length need only be a few tens of meters in length, thereby reducing latency.

The temporal soliton dragging logic gate utilizes the fiber dispersive delay line as a "lever-arm" in order to allow a weak control or signal soliton to "drag" a strong pump in time [3]. In this way, even a small spectral shift of the pump (i.e. less than a π phase shift induced by a weak signal pulse) can result in a large time delay due to non-zero group delay dispersion. Because the pump is a temporal soliton and does not broaden in time, a resolvable temporal shift can be achieved by choosing the appropriate length delay line, according to the expression

$$\Delta\omega k_0'' L > \tau_{FWHM}, \quad (1.54)$$

where the frequency shift $\Delta\omega$ is a function of the pump and signal pulses and the birefringence, and $k_0'' L$ is the group-delay dispersion. Therefore, there is a tradeoff between gain and fiber birefringence (which determine the amount of spectral shift) and gate length. If the pump were a linearly dispersive wave, broadening occurs at the same rate as the temporal delay (neglecting higher-order dispersion) and temporal resolvability can only be achieved by a resolvable spectral shift.

Temporal dragging logic must be time-shift keyed, which is still fully cascable in the optical domain (but may limit bit rate somewhat) for a clocked system, but not compatible with high-speed electronic detection. Detection can be achieved by using a trapping gate in the last stage [3] or an ultrafast optical gating mechanism, such as the Kerr gate. One problem with time-shift keyed logic is that the dragged pump is difficult to remove from the system (although it is blocked by the polarization beam splitter at the output of every second dragging gate) and may cause interference in subsequent switching stages. Note that, in contrast to the problems imposed by birefringent walkoff on the previous fiber gates, the temporal soliton trapping/dragging

¹when neglecting the nonlinear four-wave mixing effects

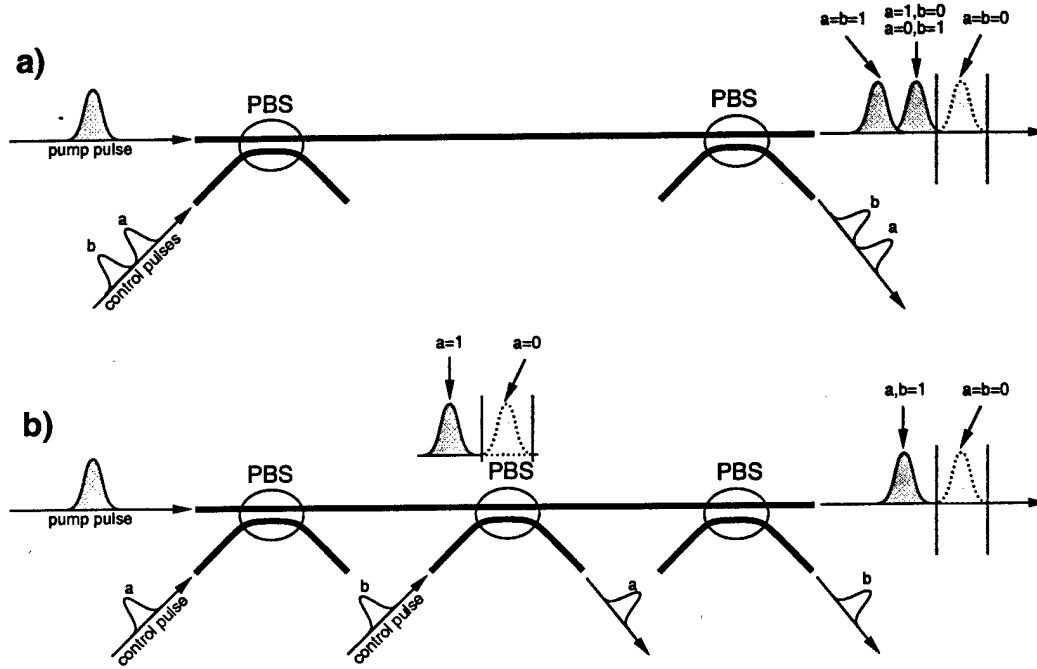


Figure 1.17: Temporal soliton dragging NOR gate geometries. Diagram a) shows a two-input soliton dragging gate which implements NOR. Diagram b) show two cascaded single-input soliton dragging gates which also implements NOR.

gates use birefringent walkoff (in combination with nonlinear cross-phase modulation) to induce the switching operation. The only constraint is that the walkoff length should be greater than about one soliton period because the bulk of the interaction occurs within that distance [66].

The first demonstration [59] of temporal soliton dragging was of a two-input, logically complete NOR gate. Two possible realizations of this logic gate are shown in Figure 1.17, where the first geometry is the one used in the experiment. In this experiment, the 300 fs pump pulse had energy of 132 pJ, while the non-soliton control pulses had 30 pJ energy, realizing a gain of 4.5 with high contrast. The fiber length was 75 meters, giving a gate latency of 360 ns. A cascading demonstration was also performed using a geometry similar to that shown in Figure 1.17 a). This demonstration suffered from the effects of Raman amplification, in which lower spectral components of a pulse are amplified at the expense of higher spectral components, resulting in asymmetric temporal broadening of the pump pulse, and may limit cascability due to the shift in mean frequency. When both control pulses are present, the output of the device is sensitive to the relative phase between the control inputs.

Using the geometry of b), a follow-on work [66] obtained switching with reduced control pulse energy and no phase sensitivity. In this demonstration, a 500 fs, 54 pJ pump pulse was used along with 6 pJ control pulses. The pulse durations were lengthened and energies were lowered to reduce the effects of Raman gain. For this geometry, the length of fiber between the coupling region for each control pulse need only be that required for the interaction to occur (about 75 meters in this case). The long dispersive delay line (350 meters) can be placed at the end in order to achieve the timing shift due to the spectral shift induced by either control. This is one instantiation of the more general time-domain chirp architecture [3]. Dragging was obtained with lower control pulse energies because each control pulse completely overlapped the pump pulse in time, resulting in a stronger interaction than in the partial overlap geometry of a).

Subsequent extensions to the time-domain chirp architecture used 30 m of moderately birefringent fiber in which the spectral shift occurred, followed by 2 km of polarization maintaining fiber as the dispersive delay line to obtain the timing shift [3]. In this case, the control pulse energy was about 1 pJ. Another modification used a short (~ 2 mm), highly nonlinear semiconductor as the nonlinear chirping element, followed by 600 m of fiber. Control energies were on the order of 100 pJ because the control and pump were slightly offset in time to achieve spectral shift upon trapping due to the lack of strong group velocity mismatch in the semiconductor.

These trapping and dragging interactions are described by the coupled, temporal NLS equations [67]

$$2ik_0 \left[\frac{\partial A_1}{\partial z} + \frac{1}{2} \Delta k_0 \frac{\partial A_1}{\partial T} \right] - k_0 k_0'' \frac{\partial^2 A_1}{\partial T^2} + 2k_0^2 \frac{n_2}{n_0} \left[|A_1|^2 + 2\Delta |A_2|^2 \right] A_1 = 0 \quad (1.55a)$$

$$2ik_0 \left[\frac{\partial A_2}{\partial z} - \frac{1}{2} \Delta k_0' \frac{\partial A_2}{\partial T} \right] - k_0 k_0'' \frac{\partial^2 A_2}{\partial T^2} + 2k_0^2 \frac{n_2}{n_0} \left[|A_2|^2 + 2\Delta |A_1|^2 \right] A_2 = 0, \quad (1.55b)$$

where $\Delta k_0'$ is the difference between the group delay coefficients, and the reduced time is defined $T = t - k_0' z$ with k_0' being the average group delay coefficient. Here, the group-delay dispersion coefficients are assumed equal and A_1 and A_2 represent the orthogonal polarizations. Note that the vector four-wave mixing terms have been neglected due to unavoidable intrinsic, stress-induced, or bend-induced birefringence which leads to accumulating phase difference between the orthogonal polarizations. Much more will be said about this fundamental system of equations in later chapters of this thesis.

1.5.2 Angular Deflection Gate based on Spatial Soliton Dragging

Spatial dragging gates [58] are one instantiation of a more general class of angular deflection logic gates studied in this thesis. The first spatial trapping gate was demonstrated experimentally [68] using spatial solitons of the same polarization (i.e. phase-dependent), but propagating at different angles. Due to cross-focusing, which is the spatial analogy to cross-phase modulation, the solitons mutually attract and under the right condition, form a trapped pair. The analogous temporal trapping interaction between two temporal solitons of different color (and hence different group velocities), has been studied analytically [64] and experimentally [63]. The direct spatial analogy to the Islam temporal trapping [62] and dragging [59] gates is the interaction between spatial solitons in the orthogonal eigenpolarization states of a uniaxial crystal [69]. In linear propagation, these eigenstates would normally walk-off in space, but nonlinear cross-focusing causes trapping or dragging to occur. A different arrangement, which is the one used in this thesis, has also been analyzed theoretically [58, 70], in which the interaction occurs between tilted orthogonally-polarized spatial solitons in linearly isotropic media. The temporal analog to this interaction is temporal solitons of different color and orthogonal polarization.

The use of solitons or solitary waves in optical logic is critical in that solitons beat the diffraction and/or dispersion limit over distances much longer than the characteristic linear lengths. Here, the main interest is in the use of lateral spatial confinement over distances larger than the linear diffraction distance. The key idea is illustrated in Figure 1.18, which shows the basic angular deflection logic gate. A pump soliton (left-hand side) propagates the length of the gate (which is assumed to be in a slab waveguide geometry) and passes through a spatial aperture at the output, forming the high output state of the device. Because the spatial soliton does not diffract, the size of the aperture can be the same as the size of the wave at the input, independent of the actual gate length.

The switching operation is performed by disturbing the propagation of the pump beam such it does not exit the spatial aperture, thus providing the low output state of the device. This can be accomplished by inducing a change in propagation angle which leads to a spatially-resolved shift at the output. If the beam propagates linearly (as shown on the right-hand side), a spatially-resolved shift can only be produced by inducing a change in the propagation angle (with a change in phase across the aperture of the beam of at least π) that is greater than twice the linear diffraction angle. In nonlinear soliton propagation, an induced angle change that is less than linearly resolvable (i.e. less than a π phase change across the spatial aperture) results in a differential spatial shift which can be integrated over non-diffracting propagation such that a spatially-resolvable shift occurs at the output. Thus, as shown in the figure, the gate length must be at least the minimum resolvable dragging distance, which depends on the spatial width of the soliton beam and the amount of angular change, much like the case of the deflection of a linear pump by a much larger angle and well within the linear diffraction length.

The angular change is produced through the nonlinear interaction between the pump and another, orthogonally-polarized beam, called the signal, which initially overlaps the pump and propagates at a non-zero relative angle. The signal beam must be strong enough to induce a nonlinear index change felt by the pump through cross-focusing. If the pump propagates linearly, then large-signal gain is not possible because the signal beam is necessarily stronger. In this case, there is no mutual nonlinear interaction and the nominal effect of the signal is to create a nonlinear prism which deflects the pump in the direction of the signal. The deflection angle depends on the relative propagation angle and intensity of the signal, with the maximum angle occurring when the pump is completely guided, or trapped, by the signal. Therefore, the angle θ must be at least the diffraction angle of the pump because the deflected pump will be nonlinearly guided by the signal, which may be a spatial soliton in the limiting case.

If the pump is a spatial soliton, then large-signal gain is possible, but the interaction is more complicated due to mutual nonlinear coupling:

$$2ik_0 \frac{\partial A_y}{\partial z} + 2ik_x \frac{\partial A_y}{\partial x} + \frac{\partial^2 A_y}{\partial x^2} + 2k_0^2 \frac{n_2}{n_0} \left[|A_y|^2 + 2\Delta |A_x|^2 \right] A_y = 0 \quad (1.56a)$$

$$2ik_0 \frac{\partial A_x}{\partial z} + \frac{\partial^2 A_x}{\partial x^2} + 2k_0^2 \frac{n_2}{n_0} \left[|A_x|^2 + 2\Delta |A_y|^2 \right] A_x = 0, \quad (1.56b)$$

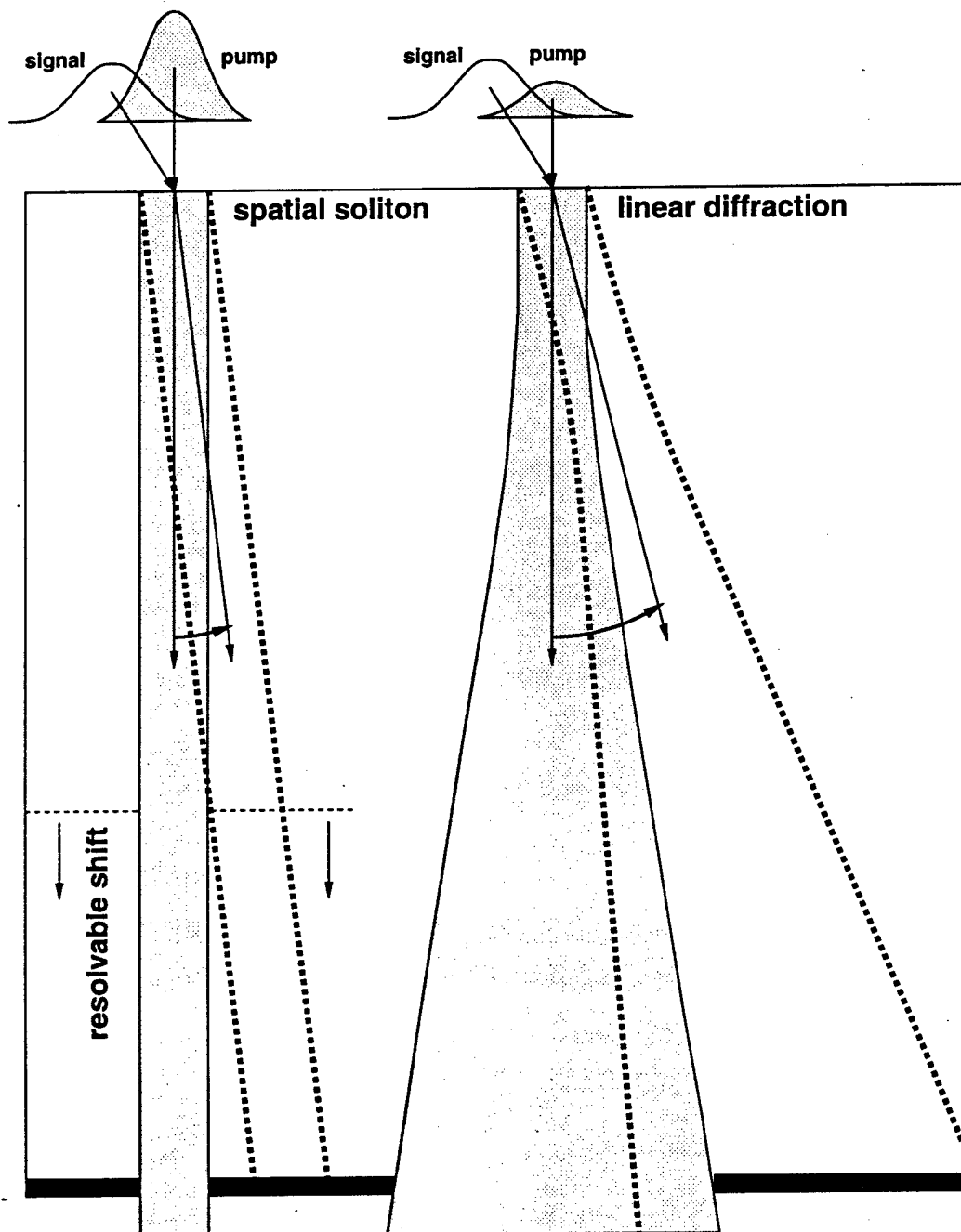


Figure 1.18: Logic gate geometry based upon the light-induced deflection of an optical "pump" beam away from a spatial aperture at the output. At one extreme, the pump can propagate nonlinearly as a spatial soliton (left), or at the other extreme, the pump can propagate linearly and diffract (right). Deflection is induced by the cross nonlinear interaction with a signal beam which is tilted with respect to the pump. The dashed contours represent the deflected pump.

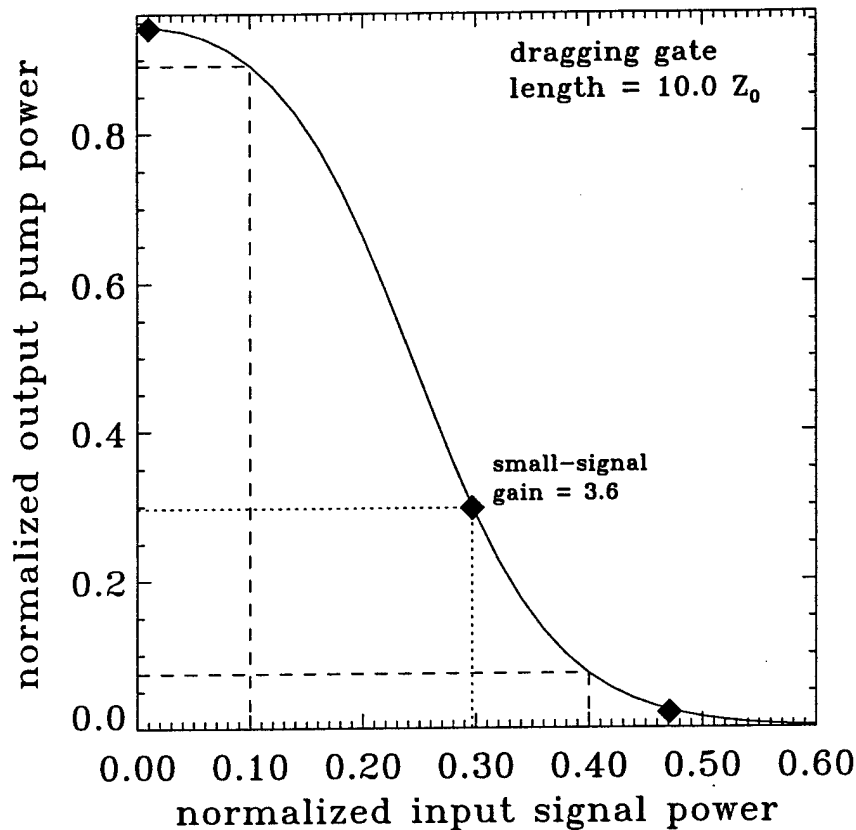


Figure 1.19: Transfer function for spatial soliton dragging logic gate demonstrating small-signal gain and saturated levels. The pump and signal are fundamental solitons of width w_0 . The signal soliton amplitude is varied to obtain the increasing input power as indicated on the horizontal scale. This transfer curve illustrates a large-signal gain of 2.0, as discussed in section 5.4.

where, for consistency with temporal interaction equations, the angle of propagation θ of the signal A_y (which is analogous to group delay) is included explicitly in the paraxial approximation such that $\theta = k_x/k_0$. Here, k_0 is the wavenumber, n_2 is the nonlinear Kerr index, and A_x is the pump envelope. This is the base system of equations used in Chapter 5 for the study of the interaction between spatial solitons. Now, during propagation, each beam affects the other through cross-focusing. If large-signal gain is realized, then the pump exerts a greater attractive force on the signal than the signal does on the pump. Nevertheless, as mentioned previously, only a small angular deviation of the pump is needed. As in the case with the linear pump, two interaction scenarios can occur: each soliton simply deflects the other, or they form a bound, orbiting pair propagating at the weighted-mean angle [58]. The former is typically referred to as dragging [59], while the latter is referred to as trapping [65]. Soliton trapping is more likely to occur when the relative propagation angle is small and the solitons are nearly the same size, while deflection, or dragging, occurs for large angles and/or large gain. Note also that the optimal interaction may not occur in complete trapping, in which case part of the signal, the "shadow" [71], may remain bound to the pump, while the rest propagates at a much larger angle as an unbound linear diffractive wave. Therefore, this interaction will be generically referred to as dragging since the high gain situation is of ultimate interest.

Figure 1.19 shows a typical input-output relation for the spatial dragging logic gate. Note that the same function would be obtained for the temporal dragging gate as well. The most important feature to note is the region of small-signal gain at the input threshold level near 0.3, surrounded by saturated levels. The transfer function shows that the dragging gate has the same operational characteristics as an nMOS inverter [26], as shown in Figure 1.1. Here, the role of the electric power supply is provided by the pump wave and the role of the gate voltage is played by the signal wave. The presence of small-signal gain at the threshold level allows the output to be driven low with very sharp switching characteristics. The threshold level is the point on the curve in which the input and output levels are the same. For an input level beyond threshold, the output of the device is switched into a valid low state; therefore, the threshold level is the minimum input signal level required to switch the output

state of the device, and determines the operating point on the transfer curve where large signal gain is unity. When small signal gain is greater than unity at the threshold level, then the input switching level can be chosen such that large signal gain and contrast greater than unity can be obtained. The transfer curve is terminated by saturated levels, showing that over switching is not possible. As a result, when operating well into the "on" and "off" regimes, small variations in the input signal level are attenuated and do not affect proper gate operation, thus providing large noise margin. A more detailed examination of soliton logic gates, based on spatial interactions, is presented in Chapters 5 and 6.

Because these gates have two inputs and only one output, they are three-terminal devices with input-output isolation. Only the undeflected pump is passed on to subsequent gates, and it is important that the pump propagate stably over the length of the gate with little change in its physical parameters. This allows for restoration of not only the logic level (because of small-signal gain), but also restoration of timing, position, polarization, color, and shape, which is crucial to cascaded operation [25]. Cascading of these gates can result in more complex logic functionality, as discussed in the next paragraph. Since the pump and signal have orthogonal polarizations, the interaction between them is phase insensitive. In the case of linear polarizations, this is not strictly true because of the presence of phase-dependent vectorial four-wave mixing terms in the nonlinear polarization. Here, it is assumed that these terms can be neglected because of waveguide birefringence, which causes each polarization to propagate with a different phase velocity. This assumption is valid when the interaction length is much greater than the birefringence beat-length.

In analogy with temporal dragging logic gate [66] of Figure 1.17 b), logically complete, n -input NOR gates can be implemented with spatial dragging in an n -stage system in which a cascaded pump is dragged to the side and blocked by the presence of a signal in any stage, thereby producing a low output, as shown in Figure 5.35 for a two-input NOR. In this implementation, the same pump is passed through two (or more) subsequent stages, so that standardized output levels (in terms of both the inverter and the multi-input NOR) may not be obtained due to absorption incurred in the additional stages. Complete logic level restoration is not strictly obtained either, because broadening of the pump in the presence of absorption also depends on total propagation distance, as discussed in section 5.3, such that the shape is not necessarily restored. However, complete logic level restoration is not necessary for a small cascade of gates (just as with electronic pass gates) and it is important to note that the pump soliton is the one that passes through multiple levels. Absorption will be the ultimate limit to the fan-in (number of stages) of this type of NOR gate. Another implementation, in which the total length of a multi-input NOR gate is held constant, so that the length of each signal input stage gets smaller with the degree of fan-in, results in both standardized output levels and complete restoration, and is discussed in more detail in sections 5.4.3 and 6.3.2.

The main benefit of the spatial dragging logic gate over the temporal dragging gate is that amplitude keyed logic is straightforward to implement while providing gain. In the spatial case, an aperture can be used to discriminate the output. For temporal dragging though, an ultrafast time gating mechanism is required to implement amplitude keyed logic instead of the more natural time-shift keyed logic [3]. A spectral filter is all that is necessary to implement amplitude keyed logic for temporal trapping, but the leverage of the dispersive delay line is lost and gain cannot be provided. Therefore, only the spatial dragging geometry has the simultaneous advantages of simple output state determination and the leverage of non-diffracting propagation which allows for large-signal gain. Additional advantages are spatial parallelism, as discussed previously, and latency. Typical spatial soliton logic gates as studied in this thesis are on the order of cm's long, while typical temporal soliton gates are on the order of 10's of m's. This results in a factor of 1000 reduction in latency. Much shorter temporal solitons can be used to reduce the length of the fiber-based gates, but higher-order temporal effects can cause problems as discussed in Chapter 6.

Spatial dragging of purely spatial solitons loses the advantages of gate-level temporal pipelining and low energies (due to complete confinement) of the temporal dragging gates, though, and spatial dragging between non-soliton pulses could yield partial switching behavior with low contrast, as shown in previous sections. That is why, in Chapter 6, this thesis studies spatial dragging of spatio-temporal solitary waves in order to retain the best features of short gate lengths, ultrafast operation and temporal pipelining, and low switching energies. The spatial dragging gate utilizing spatio-temporal solitons is perhaps the only all-optical switching and logic technology that satisfies all of the requirements for digital switching and logic and has the potential for implementation into large scale systems utilizing the high degrees of spatial and temporal parallelism available to optics.

Figure 1.20 shows the generalization of the deflection gate architecture to the use of 2-D spatio-temporal solitary waves. Utilization of fully-confined spatio-temporal waves allows for low energy operation and the additional flexibility to use spatial and/or temporal interactions. Here the focus is on spatial dragging, but temporal trapping/dragging may also occur and allow for the reduction of the requirements on allowable timing jitter between the pump and signal. Note that after the first gate, timing is restored in the system (because only the undeviated pump clock stream is passed) such that jitter no longer becomes a concern. Also, since the solitary wave does not disperse, temporal pipelining can be employed at the gate level, resulting in very high throughput which is independent of actual gate length. This architecture maps well onto the processing of a large number of independent data streams, such as ultra-fast time-division multiplexed transmission, bit-serial computation [20], and bit-serial digital signal processing [18], as discussed briefly in section 1.2.

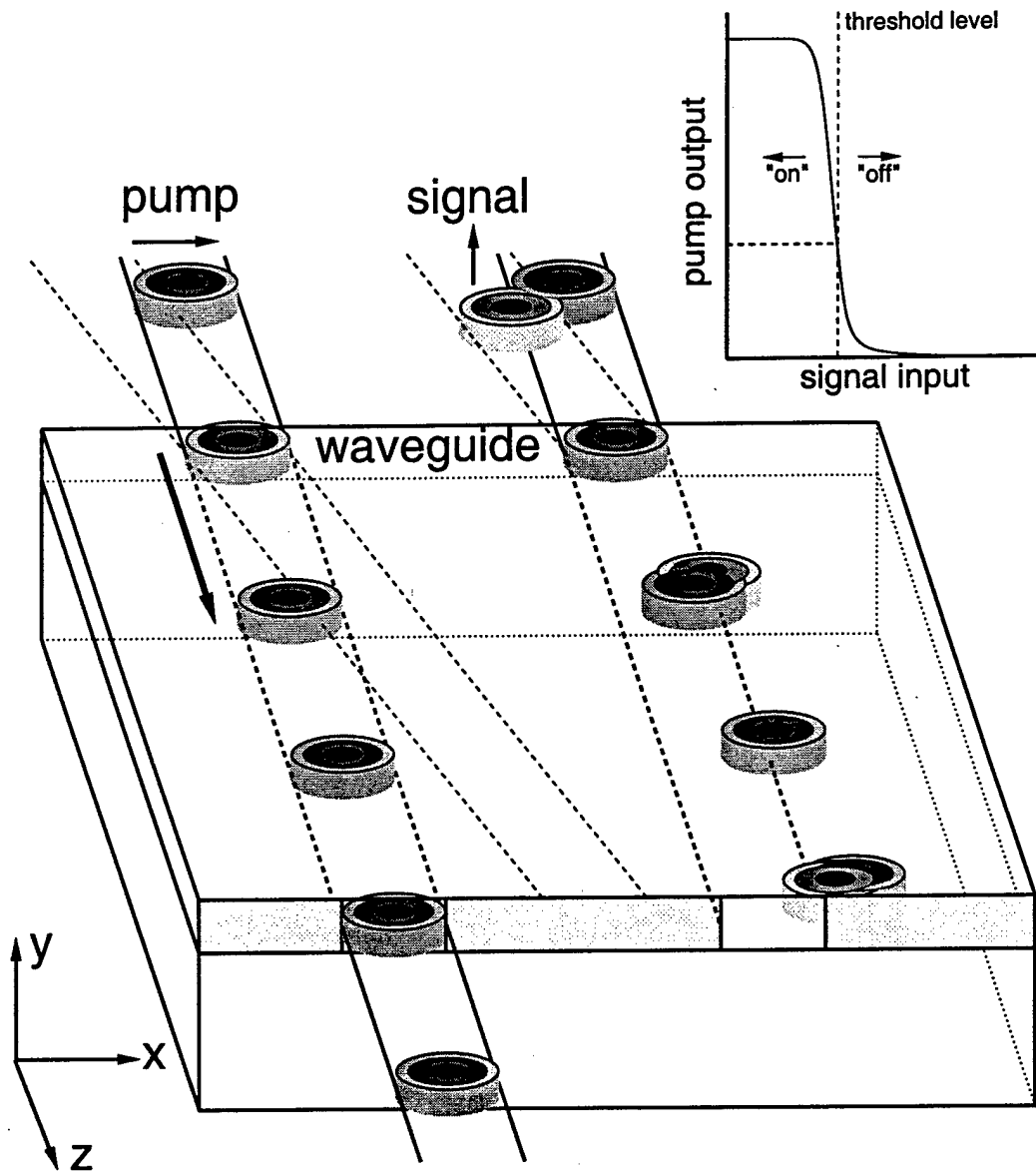


Figure 1.20: Spatio-temporal solitary wave dragging logic gate. A pump solitary wave (left) will propagate the length of the gate and pass through a spatial aperture at the output. In the presence of a tilted, orthogonally-polarized signal (right), mutual deflection will cause the pump to shift to the side and not pass through the aperture. The inset schematically shows the input-output relation for the gate. The threshold level is set at the point where the input and output are equal. This gate performs an inversion operation and can achieve large-signal gain.

It is important to make clear that dragging is an asymmetric interaction. The nonlinear waves are brought into temporal and/or spatial coincidence in linear propagation, such that they overlap at the beginning of nonlinear propagation. For a symmetric interaction, such as collision, the waves are brought into coincidence in purely nonlinear propagation where the nonlinear interaction forces are present both before and after the point of overlap. These interaction forces are symmetric in that the net change in frequency or angle is zero (the change before complete overlap is exactly compensated by the equal and opposite change after overlap), but with a small (not resolvable) residual temporal or spatial shift. Since the dragging interaction is asymmetric, the nonlinear interaction forces occur only after the point of overlap such that the frequency or angular change is not compensated. This is the key point in understanding the advantages of soliton dragging over other, symmetric, collision geometries.

Chapter 2

Optical Solitons

Outside of the communications, switching, and computing applications, optical solitary waves and optical solitons (and solitons in general) are interesting in their own right. An entire field of mathematics is devoted to the study of nonlinear integrable equations and their soliton solutions [131]. These nonlinear equations are associated with a linear scattering problem and time-evolution equation. The bound-state eigenvalues of the linear scattering equation correspond to the soliton solutions and only those solutions remain in the asymptotic limit of large time (or long distance). The nonlinear evolution of an arbitrary initial condition is described by decomposing into the linear eigenfunction basis set given by the scattering problem. These eigenvalues are then iso-spectrally evolved and transformed back to the original space using the inverse scattering transformation [73, 161, 162]. This thesis will not make direct use of the inverse scattering transform because in general the multi-dimensional vector nonlinear wave equations employed are not known to be integrable [163] and can be solved instead via more simple numerical and approximate analytical techniques.

Numerous examples of these integrable and non-integrable nonlinear wave equations exist in physics, but the examples from optics are most closely related to this thesis. There are many mechanisms of optical nonlinearity through which solitons or solitary waves can form, many of which will be discussed in section 2.1. The most ubiquitous example is the optical Kerr nonlinearity (section 2.2), which gives rise to the integrable (1+1)-D nonlinear Schrödinger (or NLS) equation, in both spatial (section 2.3) and temporal (section 2.4) forms. The multi-dimensional NLS equation (section 2.5) is not known to be integrable [163] and describes the propagation of optical waves with additional dimensions of linear and nonlinear behavior. With suitable modifications, the NLS equation serves as the basis for describing the soliton behavior necessary for the logic gates studied in this thesis.

2.1 Historical Background

A solitary wave is a nonlinear wave phenomenon that propagates without change. The most notable recorded discovery of a solitary wave occurred in 1834 when J. Scott Russell observed a solitary water wave in the Edinburgh-Glasgow canal [164, 165]. Through laboratory experiments, Russell deduced that the wave speed is proportional to its amplitude, i.e. taller waves travel faster than shorter waves. This wave phenomenon is termed a gravity wave. He also deduced that an arbitrary initial profile will asymptotically evolve into multiple solitary waves, and that two solitary waves, with the taller one overtaking the shorter one, will interact and emerge undistorted [165]. These solitary waves were later called solitons by Zabusky and Kruskal [166]. In 1871 and 1876, respectively, Boussinesq and Lord Raleigh showed that, by assuming that the length of the solitary wave is much greater the water depth, such a solitary wave has a $\text{sech}^2()$ amplitude profile. At this point it was realized that the length of the solitary wave is inversely proportional to its amplitude, such that a shorter wave is longer than a taller wave.

The governing nonlinear wave equation for shallow water waves was derived by Korteweg and deVries in 1895 [167], now known as the KdV equation:

$$\frac{\partial u}{\partial t} + 6u \frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3} = 0, \quad (2.1)$$

written in normalized form where u is the normalized amplitude of the wave as a function of position and time. In 1955, Fermi, Pasta, and Ulam [168] studied a numerical model closely related to discrete KdV, and in 1965, Zabusky and Kruskal numerically solved the KdV equation with periodic boundary conditions [166] and found that two or more KdV solitary waves do not break up upon collision. This particle-like nature led them to coin the term "soliton". In this thesis, the term solitary wave is used when it is not known that the nonlinear wave is indeed a soliton.

The first study of solitary wave phenomena in optics was the self-trapping of one- and two-dimensional filaments [169]. This was one study out of a large body of theoretical work that arose from experimental observations of beam filamentation due to self-focusing in liquids [170], bulk glasses and crystals [171, 172], and atomic gases [173]. Self-focusing was also determined responsible for anomalously low stimulated Raman scattering threshold intensities [174]. It was quickly realized in numerical studies based on the radial NLS equation [175] that two-dimensional nonlinear propagation could become unstable when diffraction does not balance self-focusing, leading to a focusing singularity. In fact, in two and three dimensions, the operating point representing the stationary solution where nonlinearity and diffraction balance is unstable, and results in the well-known critical collapse or blow-up phenomenon due to form instability [176], referred to as large-scale self-focusing [177] in two spatial dimensions. At some point in the process of focusing to a point, the beam breaks up into many filaments giving rise to small-scale self-trapping [178, 179] and is a result of modulational instability such that small transverse amplitude or phase perturbations cause a plane wave (or large scale filament) to decay into individual beams whose subsequent propagation dynamics depend on the scale of the perturbation [178].

Blow-up is an unphysical manifestation of the multi-dimensional NLS equation and indicates that the original assumptions used in the derivation of the equation break down. One mechanism of arresting this behavior is to include saturation of the nonlinear index [180, 181], resulting in stable self-trapping, thus explaining the stability of small-scale filaments. Another numerical study [179] showed the formation of small-scale filaments by following the self-focusing of a beam and subsequent breakup into an annular structure with the center portion resembling a small-scale filament. This ring structure was a result of the analysis being based on a radial nonlinear wave equation and by symmetry is the only structure possible, but will be indicative of an experimental situation in an isotropic medium in which strict radial symmetry of the initial conditions is maintained. The full picture of filamentary structure via transverse instability can only be captured with a two-dimensional analysis [178]. It was also realized that other higher-order nonlinear effects such as Raman and Brillouin scattering arrested collapse. The vector nonlinear wave equation derived in Chapter 3 includes non-paraxial terms which should also be included in the study of two-dimensional self-focusing, and will always prevent self-focusing to transverse cross-sectional areas smaller than order λ^2 .

In 1972, Zakharov and Shabat [73] published the seminal paper on the inverse scattering transform for the (1+1)-D spatial optical NLS equation. A year later, Hasegawa and Tappert [132] derived the (1+1)-D temporal NLS equation for single-mode optical fiber and proposed the use of temporal solitons as data carriers for long-haul communications. It wasn't until 1980, though, that the first temporal optical soliton was observed experimentally [182]. This experiment was made possible by the availability of low-loss optical fiber in the anomalous dispersion regime (i.e. wavelengths longer than about 1.28 μm in fused silica fiber) and the color-center laser producing ps pulses in that regime. More recent work has focused attention on the modifications of the base NLS equation required to adequately describe long distance, short-pulse propagation in fiber. These modifications include: higher-order dispersion [30], such as third- and fourth-order; optical shock [183] (also called self-steepening or the intensity-dependent group-velocity [184]); and stimulated Raman scattering [185, 186] (responsible for the soliton self-frequency shift [145, 187]). Spatial optical solitons have also been realized experimentally in liquid CS_2 [188, 189], glass [118, 120] and AlGaAs [190] slab waveguides, and in CS_2 cells using highly elliptical beams [191] to avoid critical collapse inherent in (2+1)-D nonlinear propagation [175]. Much more theoretical and experimental attention has been paid to temporal solitons, though, due to their potential for important applications, such as long-distance fiber communications.

The previous paragraphs discussed optical solitons formed through the Kerr nonlinearity, but there are other nonlinear mechanisms in optics that also allow for solitons or solitary wave behavior. It has been known since the early 1970s that a third-order like nonlinearity can be obtained by cascading two second-order nonlinearities, such as occurs with the interaction between a fundamental and a field generated through DC rectification [192] or a generated second-harmonic [193]. The former mechanism has been shown theoretically [194, 195] to be described by the Davey-Stewartson (or Benny-Roskes) equation that exhibits stabilized 2-D spatio-temporal solitons [162], but to date, these solitons have not been observed experimentally. The latter mechanism has generated a new field of study in nonlinear optics as a means of producing extremely large, ultrafast third-order nonlinear effects [196]. Second-harmonic cascading been shown theoretically [197, 198] and experimentally to support bright spatial [199] and temporal solitons.

Other nonlinear mechanisms that support solitons include: formation of temporal solitons in the Stokes pulse via stimulated inter-pulse Raman scattering (SRS) [200, 201], dark 1-D [84] and stable 2-D vortex [202] spatial solitons in self-defocusing media with thermal nonlinearity, plasma filaments in air [203], and so-called photorefractive solitons [204]. These nonlinear mechanisms do not respond on the fs time scales of the Kerr and cascaded nonlinearities and, as a result, may not be as useful for ultrafast optical switching.

2.2 Preliminaries

This thesis concentrates on bright one-dimensional spatial solitons and multi-dimensional spatio-temporal solitary waves in media exhibiting optical Kerr nonlinearity. The entire class of dark solitons are omitted from study because they have nonzero boundary conditions at infinity, implying the requirement of infinite power/energy. This boundary condition can be relaxed somewhat by embedding a dark stripe or vortex within a localized bright diffracting background. A multi-dimensional hybrid bright/dark solitary wave, or symbion, may also exist for either bright in time and dark in space (which might be important for materials with $n_2 < 0$ and normal dispersion [205]), or bright in space and dark in time (which might be important for materials with $n_2 > 0$ and normal dispersion).

In one dimension of space, a spatial soliton exists through the balance between linear diffraction and nonlinear self-focusing [73, 169]. In time, a temporal soliton exists through the balance between linear dispersion and nonlinear self-phase modulation [132]. Kerr-type (nonresonant and non-dispersive) refractive nonlinearity is responsible for both self-focusing and self-phase modulation and is the fundamental requirement for the so-called nonlinear Schrödinger (NLS) solitons in optics which are the subject of this thesis. In one dimension, NLS solitons are stable, meaning that small amplitude or phase perturbations do not upset the balance between the opposing linear and nonlinear effects.

Under special conditions on the amplitude and beam width/pulse duration, optical solitons can form when the total refractive index is well-described by the expression

$$n = n_0 + n_2 |A|^2, \quad (2.2)$$

where n_0 is the linear index of refraction at some frequency ω_0 and n_2 is the total ultrafast nonlinear index, or Kerr index, with units of cm^2/V^2 , and A is the electric field envelope with units V/cm . The definition of the nonlinear index in terms of the third-order susceptibility, given by equation 3.59, was chosen such that the above expression for the total refractive index is valid. A further generalization includes the quintic contribution to the refractive index through ultrafast $\chi^{(5)}$,

$$n = n_0 + n_2 |A|^2 + n_4^{\text{eff}} |A|^4, \quad (2.3)$$

where the effective quintic index, with units cm^4/V^4 , is defined as

$$n_4^{\text{eff}} = n_4 - \frac{n_2^2}{2n_0}, \quad (2.4)$$

ignoring the cascaded contribution. The direct quintic index n_4 is discussed in section 3.1.3. In this case, the relative dielectric constant takes the form

$$\epsilon = n^2 = n_0^2 + 2n_0 n_2 |A|^2 + 2n_0 n_4 |A|^4 + \dots \quad (2.5)$$

It is now seen that the second term in the definition of n_4^{eff} is the result of taking the square-root of the dielectric constant to obtain the refractive index. An additional contribution, produced through the process of cascading between the fundamental and third-harmonic, is derived in the multiple-scales analysis of section 3.2.

The quintic nonlinearity will be used in the multi-dimensional solitary wave case as a means for providing stability, but first 1-D spatial and temporal solitons as described by the cubic NLS equation will be discussed in detail.

2.3 1-D Spatial Optical Solitons

Spatial solitons result from the balance between linear diffraction and nonlinear self-focusing as shown in Figure 2.1. In one transverse spatial dimension, these solitons are stable such that the balance between the linear and nonlinear effects prevents small amplitude or phase perturbations from destroying the soliton. If a perturbation acts to widen the soliton, nonlinear self-focusing overpowers diffraction to restore balance. In the opposite case, if a perturbation narrows the soliton, linear diffraction overtakes self-focusing.

The analytical study of spatial solitons begins with the scalar Helmholtz equation which describes the propagation of a monochromatic beam in weakly inhomogeneous media

$$\nabla^2 \bar{A} + k^2 \bar{A} = 0 \quad (2.6)$$

where $E = \frac{1}{2} \bar{A}(x, y, z) e^{-i\omega_0 t} + \text{cc}$ is the time-harmonic electric field, $k = \omega_0 n(\mathbf{r}, \omega_0)/c$ is allowed to have weak spatial variation, and the over bar indicates the presence of a fast phase variation due to propagation. Assuming for the moment that k has no spatial variation, equation 2.6 can be transformed into the spatial frequency, or Fourier, domain

$$[k_x^2 + k_y^2 + k_z^2] \bar{A} = k_0^2 \bar{A}, \quad (2.7)$$

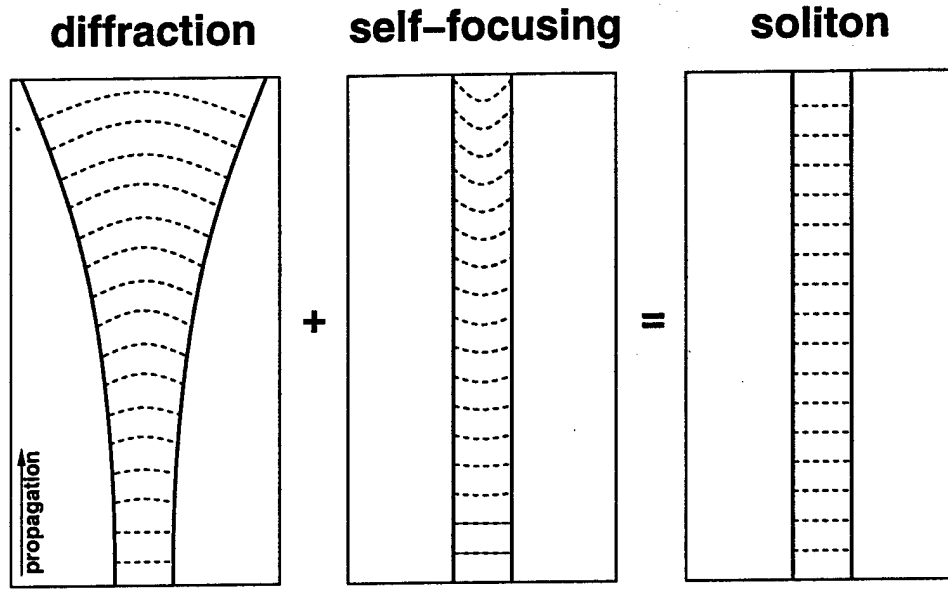


Figure 2.1: Illustration of the balance between linear diffraction and nonlinear self-focusing in the formation of a spatial optical soliton. The $1/e^2$ beam radius is plotted along with the dashed lines which represent the phase fronts.

where k_x , k_y and k_z are spatial frequency variables representing the Cartesian components of the optical wave vector. The solution to this equation is

$$\sqrt{k_x^2 + k_y^2 + k_z^2} = k_0, \quad (2.8)$$

which describes the momentum sphere of isotropic propagation and prescribes one spatial frequency component when the other two are known. This sets up an initial-value problem for propagation, such that when the electric field is known on a plane for example, guaranteeing that two of the three components of momentum are known, propagation to another parallel plane at an arbitrary distance is then fully determined by specifying the third component using equation 2.8. If z is the direction of propagation, then the initial-value problem has the solution in the Fourier domain

$$\tilde{A}(k_x, k_y, z) = e^{\pm i \sqrt{k_0^2 - k_x^2 - k_y^2} z} \tilde{A}(k_x, k_y, 0), \quad (2.9)$$

where $\tilde{A}(k_x, k_y, 0)$ is known by the Fourier transform of the initial field on the input plane. The field distribution $\tilde{A}(x, y, z)$ at the observation plane is calculated by the inverse Fourier transform of $\tilde{A}(k_x, k_y, z)$. Linear diffraction is described in more detail in section 2.3.2 and Chapter 4.

In media of isotropic symmetry class with Kerr-type nonlinear index n_2 , the inhomogeneous wave number is

$$k^2(|A|) = k_0^2 + 2k_0^2 \frac{n_2}{n_0} |A|^2 \quad (2.10)$$

where the quintic nonlinear index is neglected. Substituting into equation 2.6 results in

$$\nabla^2 \tilde{A} + k_0^2 \tilde{A} + 2k_0^2 \frac{n_2}{n_0} |A|^2 \tilde{A} = 0. \quad (2.11)$$

This equation describes the (2+1)-D spatial propagation of a beam in a weakly inhomogeneous, weakly nonlinear medium. Since the 1-D spatial soliton is of present interest, it is assumed that the beam is confined in the y dimension by a slab waveguide. As a result, assuming that the linearly guided and nonlinear soliton envelopes are separable, i.e.

$$\tilde{A}(x, y, z) = \Phi(y) \tilde{A}(x, z), \quad (2.12)$$

where $\Phi(y)$ represents the guided profile and is real and of unity magnitude, the nonlinear Helmholtz equation 2.11 reduces to

$$\frac{\partial^2 \tilde{A}}{\partial z^2} + \frac{\partial^2 \tilde{A}}{\partial x^2} + k_0^2 \tilde{A} + 2k_0^2 \frac{n_2}{n_0} |A|^2 \tilde{A} = 0 \quad (2.13)$$

and is the starting point for the following analysis of 1-D spatial solitons.

Note that the effects of waveguide confinement have been formally neglected in equation 2.13. This omission can be remedied by using an effective wavenumber β^{eff} and nonlinear coefficient n_2^{eff} determined by the appropriate modal overlap integrals, but the analysis in the next two sections neglects the effects of transverse waveguide confinement for simplicity. The fully confined case of 2-D spatio-temporal solitary waves in a slab waveguide is discussed in full in Appendix C.

2.3.1 The Non-Paraxial Fundamental Soliton

A soliton is by definition a stationary solution, therefore, using the field amplitude ansatz

$$\bar{A}(x, z) = A(x)e^{i\beta z} \quad (2.14)$$

in equation 2.13, the stationary transverse envelope becomes the eigenfunction of the ordinary nonlinear differential equation

$$\frac{d^2 A(x)}{dx^2} + \left[k_0^2 - \beta^2 + 2k_0^2 \frac{n_2}{n_0} |A(x)|^2 \right] A(x) = 0, \quad (2.15)$$

where β^2 is the eigenvalue and β represents the total wavenumber and contains both linear and nonlinear contributions.

The transverse envelope is assumed to be of the form [169]

$$A(x) = A_0 \operatorname{sech} \left(\frac{x}{w_0} \right), \quad (2.16)$$

where A_0 is the real amplitude and w_0 is a measure of the beam width. Since the soliton is non-diffracting, there is no transverse phase variation. The stationary solution is given by

$$A_0 = \frac{1}{k_0 w_0} \sqrt{\frac{n_0}{n_2}} \quad (2.17)$$

$$\beta^2 = k_0^2 + \frac{1}{w_0^2} = k_0^2 \left[1 + \frac{n_2 A_0^2}{n_0} \right], \quad (2.18)$$

with the total wavenumber $\beta \approx k_0 + k_f n_2 A_0^2 / 2$ when the induced nonlinearity is small.

Since equation 2.13 is invariant under the unitary transformation

$$\begin{bmatrix} x' \\ z' \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}, \quad (2.19)$$

a more general stationary solution is given by

$$A(x', z') = A_0 \operatorname{sech} \left[\frac{\cos \theta x' - \sin \theta z'}{w_0} \right] e^{i\beta [\cos \theta z' + \sin \theta x']}. \quad (2.20)$$

Equation 2.20 is a fully linear and nonlinear non-paraxial solution when the optical field polarization is in the y direction such that vectorial effects are absent. Unfortunately, it is mathematically and numerically difficult to propagate this (or other) solution using the evolution equation 2.13; therefore, a common practice is to reduce this full Helmholtz equation to a parabolic one so that robust numerical techniques based on finite-difference or split-step methods, and analytic techniques such as the scattering/inverse-scattering transform, can be used. This procedure is carried out formally in Chapter 3 for the fully vectorial (3+1)-D and (2+1)-D evolution equations which uncovers the linear and nonlinear physics buried in the second-order equation 2.6, but done in an ad-hoc manner in the next section for the (1+1)-D case.

2.3.2 The Paraxial Fundamental Soliton

Paraxial propagation is described about some mean direction of propagation. For simplicity, choosing z to be this direction, the electric field envelope is assumed to be of the form

$$\bar{A}(x, z) = A(x, z)e^{ik_0 z}, \quad (2.21)$$

where $k_0 = \omega_0 n_0 / c$ is the bulk propagation constant or wave number. The fast phase due to propagation is explicitly removed from the transverse envelope so that the slowly-varying amplitude (SVA) and paraxial approximations can be made. Substituting this expression into the nonlinear Helmholtz equation 2.13, and making the approximation

$$|\partial^2 A / \partial z^2| = \frac{1}{2k_0} \left| \frac{\partial^3 A}{\partial z^3} + \frac{\partial^2 A}{\partial z \partial x^2} + 2k_0^2 \frac{n_2}{n_0} \frac{\partial |A|^2 A}{\partial z} \right| \ll |2k_0 \partial A / \partial z|, \quad (2.22)$$

results in the (1+1)-D spatial NLS equation

$$2ik_0 \frac{\partial A}{\partial z} + \frac{\partial^2 A}{\partial x^2} + 2k_0^2 \frac{n_2}{n_0} |A|^2 A = 0. \quad (2.23)$$

The first two terms in equation 2.23 simply describe linear diffraction in the paraxial approximation which is seen by taking a transverse spatial Fourier transform

$$\frac{\partial \tilde{A}}{\partial z} = -i \frac{k_x^2}{2k_0} \tilde{A} \Rightarrow \tilde{A}(z) = \tilde{A}(0) e^{-i[k_x^2 / 2k_0]z}, \quad (2.24)$$

and reduces the z -directed accumulated phase from the maximum value k_0 to $k_0 [1 - k_x^2 / 2k_0^2]$ due to the off-axis projection of the phase-velocity onto the mean direction of propagation z . This expression only allows for propagation in one direction (along z or very near to it) and is the result of the using the slowly-varying amplitude approximation which reduces the Helmholtz equation to a uni-directional wave equation. The SVAA is valid when backscattered radiation from linear or nonlinear inhomogeneity can be neglected to a given order of approximation, as discussed in Chapter 3.

Figure 2.2 illustrates diffraction in spatial frequency space. For full (2+1)-D spatial diffraction in isotropic space, the surface of allowed k vectors is a sphere. Diffraction in (1+1)-D as shown in the figure is represented by slicing the sphere with a plane passing through the origin, resulting in a circle. The paraxial approximation represents this circle as a parabola and is valid for small transverse spatial frequencies (small angular bandwidths). Linear propagation is described by appropriately phasing each transverse spatial frequency component $\tilde{A}(k_x)$ with $\exp(ik_z z)$, where $k_z = \sqrt{k_0^2 - k_x^2}$ in the non-paraxial case and $k_z \approx k_0 - k_x^2 / 2k_0$ in the paraxial case.

Figure 2.3 shows the diffraction of an initial $\text{sech}(x/w_0)$ profile beam over 5 confocal distances. Here the confocal distance $Z_0 = \pi^2 w_0^2 / \lambda$ is twice the Rayleigh range, which represents the distance over which the FWHM intensity pattern increases by $\sqrt{2}^1$. Note that the intensity FWHM is $1.7627 w_0$ [30].

Neglecting the diffraction term in equation 2.23 results in the nonlinear equation

$$i \frac{\partial A}{\partial z} + k_f n_2 |A|^2 A = 0, \quad (2.25)$$

where $k_f = \omega_0 / c$ is the free-space wavenumber. The solution to this equation is written in the real space domain as

$$A(x, z) = e^{i\phi^{NL}(x, z)} A(x, 0) \quad (2.26)$$

with the nonlinear phase accumulation

$$\phi^{NL}(x, z) = k_f n_2 \int_0^z |A(x, z')|^2 dz'. \quad (2.27)$$

For small $z = \Delta z$, the nonlinear phase is approximated by the expression

$$\phi^{NL}(x, \Delta z) \approx k_f n_2 |A(x, \Delta z/2)|^2 \Delta z \quad (2.28)$$

which is second-order accurate in Δz . This nonlinearly induced phase is illustrated in Figure 2.4 for propagation distance Δz and peak amplitude chosen such that the maximum nonlinear phase is 4 radians.

The spatial frequency space is the natural domain in which to solve the linear diffraction problem, while real space is the natural domain in which to solve the nonlinear problem (ignoring diffraction). This insight will be used in the development of the split-step Fourier method outlined in Chapter 4, which efficiently implements nonlinear propagation by switching between the two domains for linear diffraction and nonlinear refraction.

¹Unlike the case for Gaussian beams, in the case of a $\text{sech}()$, this distance is not the same as that obtained by interpolating the far-field diffraction pattern, which results in a confocal distance $1.605\pi^2 w_0^2 / \lambda$.

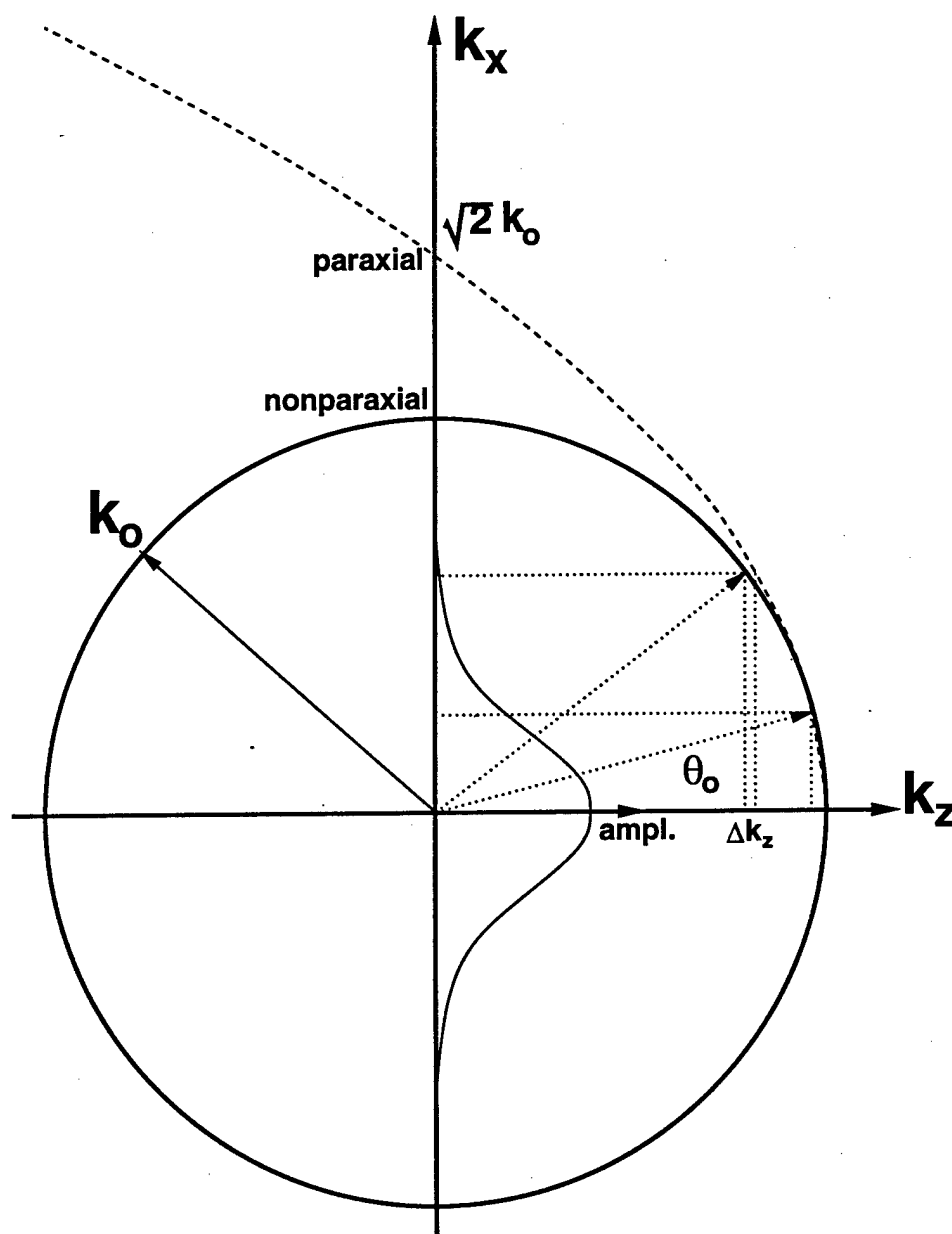


Figure 2.2: k-space diagram graphically showing the z-projected wavenumber k_z as a function of the transverse wavenumber k_x . The paraxial approximation is illustrated by the dashed curve. The transverse spatial frequency spectrum of a beam is shown in order to illustrate propagation by associating k_z with every k_x of the spectrum. The paraxial approximation is good for small k_x but breaks down for larger k_x where the associated value for k_z deviates significantly from the exact result.

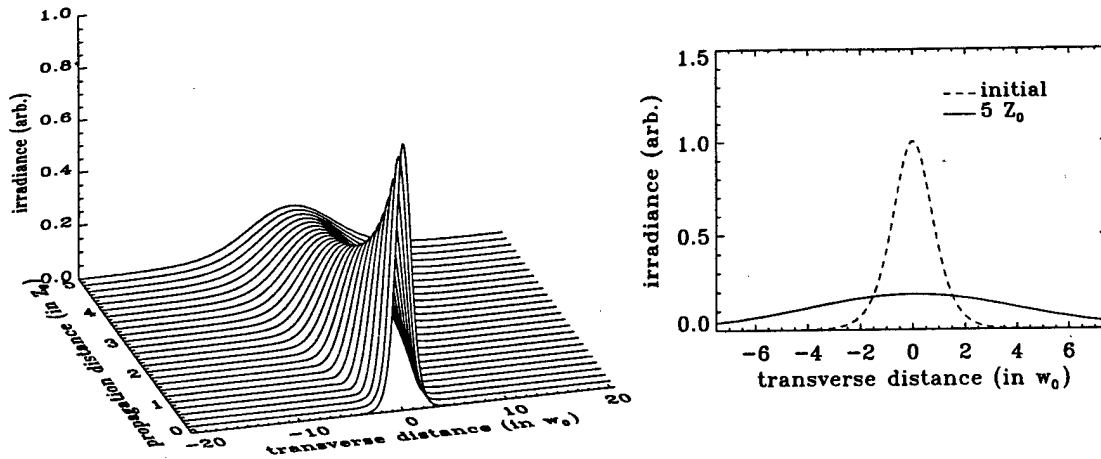


Figure 2.3: Diffraction of a beam of initial $\text{sech}(x/w_0)$ profile, and intensity profile $\text{sech}^2(x/w_0)$, where $w_0 = 14.2 \mu\text{m}$ and intensity FWHM $25 \mu\text{m}$. After propagating $5 Z_0$, where $Z_0 = \pi^2 w_0^2 / \lambda = 1.85 \text{ mm}$ is the confocal distance, the intensity FWHM broadens to $140 \mu\text{m}$.

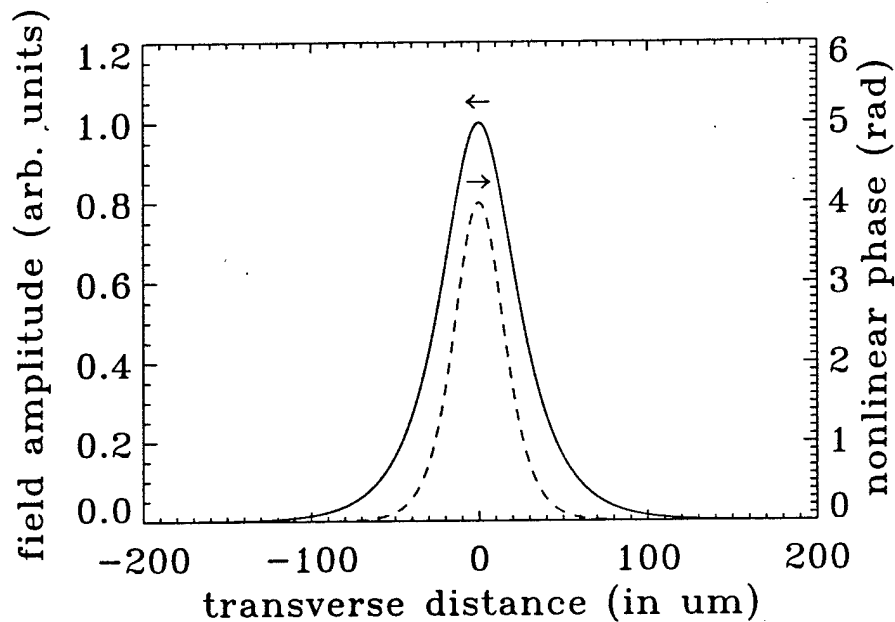


Figure 2.4: The nonlinear phase (dashed curve) induced by an electric field envelope (solid curve). The induced phase is proportional to the local intensity and in a material with positive Kerr nonlinearity, tends to reduce the phase velocity of the central portion of the beam resulting in a focusing, or lensing, effect.

Now, looking for a paraxial soliton solution, the stationary ansatz 2.14 is substituted into equation 2.23 with the result

$$\frac{d^2 A(x)}{dx^2} + 2k_0 \left[-\beta + k_f n_2 |A(x)|^2 \right] A(x) = 0. \quad (2.29)$$

Assuming the fundamental soliton form of equation 2.16, the amplitude and phase become

$$A_0 = \frac{1}{k_0 w_0} \sqrt{\frac{n_0}{n_2}} \quad (2.30)$$

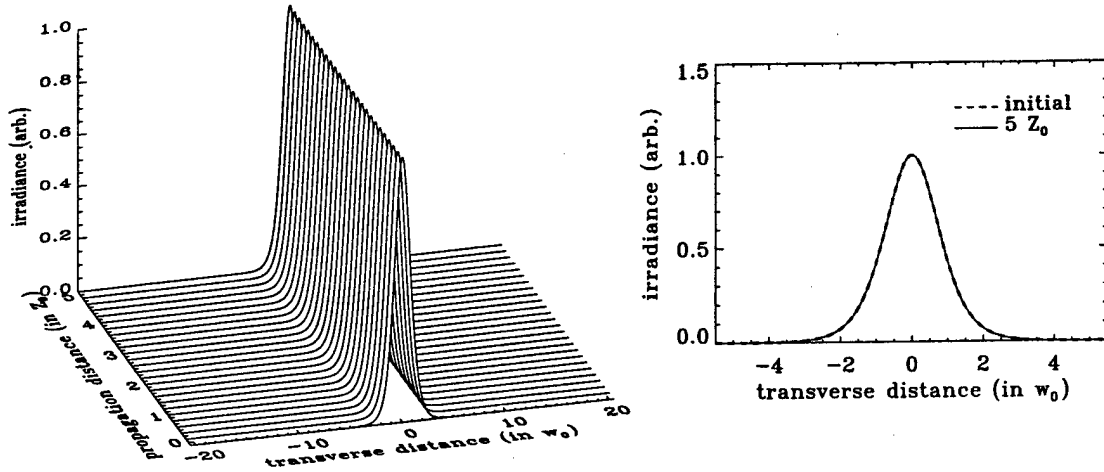


Figure 2.5: Spatial soliton beam of initial $\text{sech}(x/w_0)$ amplitude profile, where $w_0 = 14.2 \mu\text{m}$ and intensity FWHM $25 \mu\text{m}$. Even after propagating $5 Z_0$, where $Z_0 = 1.85 \text{ mm}$, the beam width and position are unchanged.

$$\beta = 1/2k_0w_0^2 = k_f n_2 A_0^2/2, \quad (2.31)$$

where the amplitude is the same as in the non-paraxial solution and β represents the slow phase nonlinear correction to the fast phase linear wavenumber k_0 . The total wavenumber $k_0 + \beta$ is the same as that obtained in the non-paraxial case given by equation 2.18 when $w_0 \gg \lambda$. The nonlinear correction to the wave number β is half that obtained for a plane wave of amplitude A_0 , and simply reflects the fact that the soliton does not have constant amplitude.

The full paraxial solution is written

$$\bar{A}(x, y, z) = \frac{1}{k_0 w_0} \sqrt{\frac{n_0}{n_2}} \Phi(y) \text{sech}\left(\frac{x}{w_0}\right) e^{i[k_0 + 1/2k_0w_0^2]z}, \quad (2.32)$$

where the linearly guided profile $\Phi(y)$ and linear wavenumber k_0 are added for completeness. Spatial soliton propagation using the split-step numerical scheme of section 4.2.1 is shown in Figure 2.5 with $w_0 = 14.2 \mu\text{m}$. After a propagation distance of $5 Z_0$, the soliton is unchanged, thereby beating the limits imposed by linear diffraction as shown in Figure 2.3, which would result in spatial broadening by a factor of 5.6.

Defining the optical intensity with units W/cm^2

$$I(x, y, z) = \frac{\epsilon_0 c n}{2} |A(x, y, z)|^2 = \frac{1}{2\eta} |A(x, y, z)|^2, \quad (2.33)$$

and neglecting the nonlinear contribution to the index and impedance, the intensity of the fundamental soliton is written

$$I(x, y, z) = \frac{\epsilon_0 c}{2k_f^2 w_0^2 n_2} \Phi^2(y) \text{sech}^2\left(\frac{x}{w_0}\right), \quad (2.34)$$

for n_2 in units of V^2/cm^2 . The power of the soliton is

$$P(z) = \int \int I(x, y, z) dx dy = \frac{\epsilon_0 c}{k_f^2 w_0 n_2} \int \Phi^2(y) dy \quad (2.35)$$

and has units of W . Note that $\int \text{sech}^2(x/w_0) dy = 2w_0$.

From these expressions, it is clear that the peak intensity of the soliton increases inversely proportionally to the square of the width and the power of the soliton increases inversely proportionally to the width, so that a narrow soliton requires more power to launch and propagate than a broader one.

A further generalization of the paraxial soliton solution is given by

$$A(x, z) = A_0 \text{sech}\left(\frac{x - \theta z}{w_0}\right) e^{i[\delta k_x x + \beta z]}. \quad (2.36)$$

Substituting this expression into the NLS equation 2.23 results in

$$A_0 = \frac{1}{k_0 w_0} \sqrt{\frac{n_0}{n_2}} \quad (2.37)$$

$$\theta = \delta k_x / k_0 \quad (2.38)$$

$$\beta = \frac{1}{2k_0 w_0^2} - \frac{\delta k_x^2}{2k_0}, \quad (2.39)$$

which is the paraxial version of equation 2.20. Notice that, like the non-paraxial soliton solution, the NLS soliton solution has two free parameters: the width w_0 and angle of propagation θ .

2.4 1-D Temporal Optical Solitons

Temporal optical solitons result from the balance between linear dispersion and nonlinear self-phase modulation (SPM). In order to arrive at this one-dimensional system, the temporal soliton is confined transversely in an optical fiber. Even though the nonlinearity of silica glass is very small, optical fiber is an ideal system in which to perform experimental studies because of the very low linear loss ~ 0.2 dB/km ($\sim 5 \times 10^{-7}$ cm $^{-1}$) and long interaction lengths [30]. Because a temporal soliton is inherently a one-dimensional nonlinear phenomenon described by the NLS equation, stability is guaranteed as it is in the one-dimensional spatial case.

The scalar wave equation describing temporal propagation in fiber is

$$\frac{\partial^2 \bar{A}}{\partial z^2} + k^2 \left(\omega_0 + i \frac{\partial}{\partial t} \right) \bar{A} + 2 \frac{n_0 n_2}{c^2} \left[\omega_0 + i \frac{\partial}{\partial t} \right]^2 |\bar{A}|^2 \bar{A} = 0, \quad (2.40)$$

where the electric field $E = \frac{1}{2} \bar{A}(x, y, z, t) e^{-i\omega_0 t} + \text{cc}$ and Raman scattering has been neglected. The wave number function of the operator $\frac{\partial}{\partial t}$ is interpreted by its Taylor expansion [194], as shown in equation 2.44. Here, parenthesis enclose function arguments, while braces and brackets are used for grouping. The effects of Raman scattering are considered in section 2.4.3. Equation 2.40 describes the evolution of the pulse envelope, therefore soliton solutions are termed envelope solitons, in contrast to field soliton solutions of other nonlinear wave equations such as KdV.

In the following analysis it is assumed that \bar{A} has the separable form $\bar{A} = \Phi(x, y) \bar{A}(z, t)$, where Φ is the transversely guided profile of unity magnitude. The effects of transverse guidance will be neglected for simplicity, as already done in equation 2.40. In Appendix C, the derivation in the case of spatio-temporal propagation in a slab waveguide is presented in full.

The linear part of equation 2.40 has the simple solution in the temporal frequency domain

$$\bar{A}(\Delta\omega, z) = e^{\pm ik(\omega_0 + \Delta\omega)z} \bar{A}(\Delta\omega, 0). \quad (2.41)$$

The term $\Delta\omega$ describes some (small) frequency variation about the central frequency ω_0 . Similar to the case for linear diffraction, linear dispersive propagation is fully described by phasing each temporal frequency component $\bar{A}(\omega_0 + \Delta\omega)$ using the appropriate dispersive wave number $k(\omega_0 + \Delta\omega)$.

The variation in wavenumber k with wavelength is shown in Figure 2.6 for bulk silica glass, where waveguide dispersion is neglected. The linear and nonlinear optical properties of silica are discussed in more detail in Appendix D. Also shown is the wavelength dependence of the group delay coefficient, which is related to the wavenumber by $k' = \partial k / \partial \omega$, and can be evaluated in two forms

$$k'(\omega) = \frac{1}{c} \left[n(\omega) + \frac{\omega}{c} \frac{\partial n(\omega)}{\partial \omega} \right] \quad (2.42)$$

$$k'(\lambda) = \frac{1}{c} \left[n(\lambda) - \lambda \frac{\partial n(\lambda)}{\partial \lambda} \right]. \quad (2.43)$$

It is clear from the figure that when $\lambda > 1.28 \mu\text{m}$, shorter wavelengths are delayed less than longer wavelengths, meaning that shorter wavelengths (higher frequencies) travel with a greater group velocity than longer wavelengths. This is the anomalous dispersion regime and is the regime in which bright temporal solitons can form with positive Kerr nonlinearity.

Much effort has been placed on using waveguide dispersion of fiber to tailor the overall dispersion characteristics. One example of this is dispersion-shifted fiber, in which the zero value of group-delay dispersion is shifted to longer wavelengths, particularly near $1.55 \mu\text{m}$ where the fiber loss is minimum. Double or quadruple clad fiber has flattened dispersion characteristics where the GDD is small and nearly uniform over a 25 nm range [30]. Both of these examples are used mainly for linear propagation where dispersion is not compensated by SPM. Appendix D briefly covers the effects of waveguide parameters on the overall linear dispersion relation.

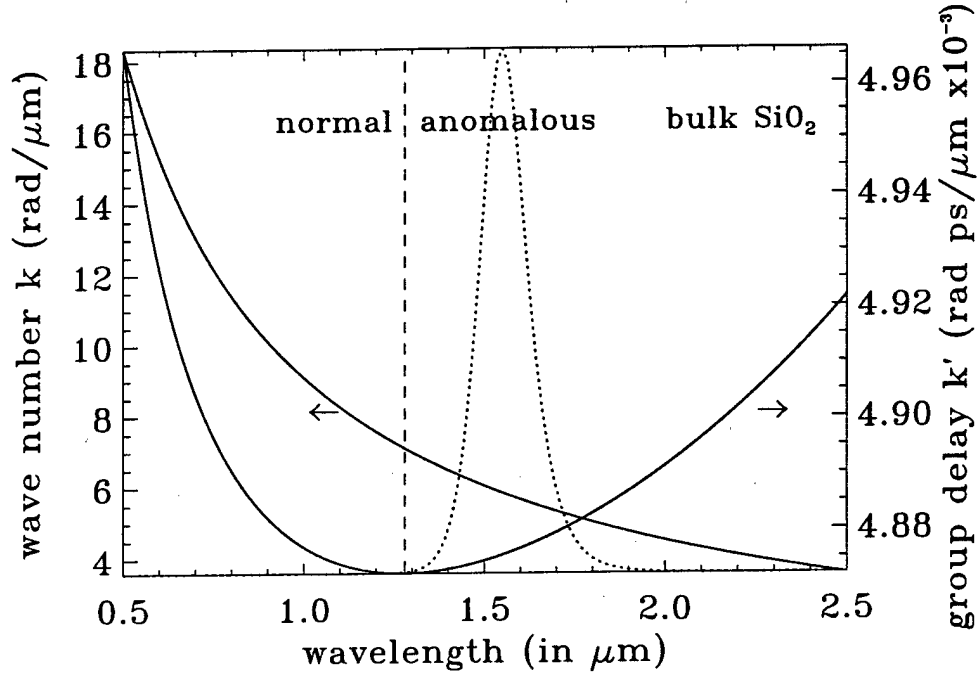


Figure 2.6: Wave number and group delay versus wavelength for bulk silica. The dashed vertical line separates the normal (NGDD) and anomalous (AGDD) dispersion regimes. The spectrum of a 16.5 fs intensity FWHM sech() pulse with time-bandwidth product 0.316 is also shown. Linear dispersive propagation occurs by phasing each temporal frequency component using the appropriate wave number.

2.4.1 The Non-SVEA Fundamental Soliton

It is not known that equation 2.40 has a general solution. There are solutions for specific material parameter values when the full dispersion relation $k^2(\omega)$ is truncated at a given order. This truncation amounts to making the slowly-varying envelope approximation (SVEA), although the standard SVEA keeps terms only up to second-order in time derivatives and results in the temporal NLS equation.

An analytic solution to equation 2.40 has been obtained [206] keeping linear dispersion terms up to fourth-order in time derivatives and all nonlinear terms, excluding Raman scattering. To fourth order, the linear dispersion is

$$k^2 \left(\omega_0 + i \frac{\partial}{\partial t} \right) \approx k_0^2 + 2ik_0 k_0' \frac{\partial}{\partial t} - [k_0'^2 + k_0 k_0''] \frac{\partial^2}{\partial t^2} - \frac{i}{3} [3k_0' k_0'' + k_0 k_0'''] \frac{\partial^3}{\partial t^3} + \frac{1}{12} [3k_0'^2 + 4k_0' k_0''' + k_0 k_0''''] \frac{\partial^4}{\partial t^4}. \quad (2.44)$$

With this approximation, equation 2.40 can be rewritten

$$\frac{\partial^2 \bar{A}}{\partial z^2} + k_0^2 \bar{A} + 2ik_0 k_0' \frac{\partial \bar{A}}{\partial t} - [k_0'^2 + k_0 k_0''] \frac{\partial^2 \bar{A}}{\partial t^2} - \frac{i}{3} [3k_0' k_0'' + k_0 k_0'''] \frac{\partial^3 \bar{A}}{\partial t^3} + \frac{1}{12} [3k_0'^2 + 4k_0' k_0''' + k_0 k_0''''] \frac{\partial^4 \bar{A}}{\partial t^4} + 2n_0 n_2 \left[k_f^2 + \frac{2ik_f}{c} \frac{\partial}{\partial t} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] |\bar{A}|^2 \bar{A} = 0. \quad (2.45)$$

Notice that the SVAA has not been made, even though the SVEA has been made to fourth-order in time derivatives. Equation 2.45 will still be referred to as a non-SVEA equation because it includes two orders of time derivatives beyond the standard SVEA equation.

A solitary wave solution to equation 2.45 is given by [206]

$$\bar{A}(z, t) = A_0 \operatorname{sech} \left(\frac{t - z/v}{\tau_0} \right) e^{i[\beta z - \delta \omega t]}, \quad (2.46)$$

where the expressions for the quantities A_0 , v , τ_0 , β and $\delta\omega$ can be found in the reference. One important property of the solution is that all of the parameters are determined by the material constants; there are no free parameters so only one solitary wave profile can be a self-consistent solution. This is unlike the NLS case in which there are two free parameters - width/duration (or amplitude) and propagation angle/frequency [117]. A condition on the existence of the solution 2.46 is that the material dispersion parameters k_0'' , k_0''' and k_0'''' are negative, which can be accomplished using quadruple-clad fiber for example [207].

Another property of this solution is that, like the non-paraxial spatial soliton solution of the previous section, there is no backscatter of radiation from the induced inhomogeneity due to the nonlinear index. This is because the wave is a stationary state and does not change with propagation, therefore there is no energy loss due to scattering. The stability properties of this non-SVEA solution have not been studied to date, and it should be pointed out that experimental realization may be difficult due to the restrictive nature of the solution.

2.4.2 The SVEA Fundamental Soliton

Writing the field amplitude in terms of a fast propagation phase and slowly-varying envelope,

$$\tilde{A}(z, t) = A(z, t)e^{ik_0 z}, \quad (2.47)$$

and substituting into equation 2.40, the (1+1)-dimensional scalar nonlinear wave equation usually considered for soliton propagation in optical fiber [132] is obtained

$$2ik_0 \left[\frac{\partial A}{\partial z} + k_0' \frac{\partial A}{\partial t} \right] - k_0 k_0'' \frac{\partial^2 A}{\partial t^2} + 2k_0^2 \frac{n_2}{n_0} |A|^2 A = 0, \quad (2.48)$$

where the slowly-varying amplitude and envelope approximations have been made. The quantity in brackets simply indicates that the wave nominally moves at the group velocity $v_g = 1/k_0'$. This quantity can be simplified by making the transformation into the reduced time coordinates

$$\begin{aligned} T &= t - k_0' z \\ Z &= z, \end{aligned} \quad (2.49)$$

and noting that

$$\frac{\partial}{\partial t} = \frac{\partial T}{\partial t} \frac{\partial}{\partial T} + \frac{\partial Z}{\partial t} \frac{\partial}{\partial Z} = \frac{\partial}{\partial T} \quad (2.50)$$

$$\frac{\partial}{\partial z} = \frac{\partial T}{\partial z} \frac{\partial}{\partial T} + \frac{\partial Z}{\partial z} \frac{\partial}{\partial Z} = -k_0' \frac{\partial}{\partial T} + \frac{\partial}{\partial Z}. \quad (2.51)$$

With these substitutions, equation 2.48 can be written

$$2ik_0 \frac{\partial A}{\partial Z} - k_0 k_0'' \frac{\partial^2 A}{\partial T^2} + 2k_0^2 \frac{n_2}{n_0} |A|^2 A = 0, \quad (2.52)$$

which is the temporal nonlinear Schrödinger (NLS) equation. Apart from a scaling factor, this equation is the same as the (1+1)-D spatial NLS equation. For bright soliton solutions analogous to equation 2.32, it is clear that $k_0'' < 0$, meaning that the wave propagates in the anomalous dispersion regime.

By using the transformation 2.49, the coordinate system is changed such that the pulse moving with the group velocity $1/k_0'$ is stationary on the reduced time axis T , instead of moving on the time axis t at the group velocity. This results in the leading-order frequency-dependent correction to the wavenumber k_0 being proportional to the group-delay dispersion coefficient k_0'' , or

$$k(\omega) \approx k_0 + \frac{1}{2} \Delta\omega^2 k_0'', \quad (2.53)$$

where $k_0 = k(\omega_0)$ and $\Delta\omega = \omega - \omega_0$.

The linear part of equation 2.52 has a simple solution in the temporal frequency domain

$$\tilde{A}(\Delta\omega, z) = e^{i[k_0'' \Delta\omega^2 / 2]z} \tilde{A}(\Delta\omega, 0). \quad (2.54)$$

Analogous to paraxial linear spatial diffraction, linear dispersion is described by phasing each temporal frequency component with the appropriate quadratic phase factor $e^{ik_0'' \Delta\omega^2 z / 2}$ in this case where SVEA was made such that higher-order dispersion is

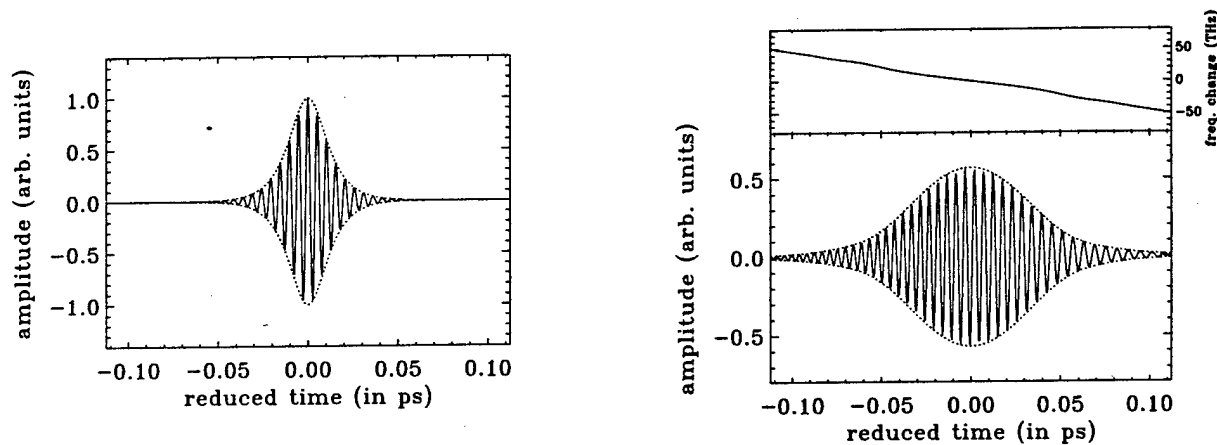


Figure 2.7: Effect of negative group-delay dispersion on a 16.5 fs quasi-monochromatic wavepacket with $\lambda_f = 1.55 \mu\text{m}$. The initial sech() packet is shown on the left and the dispersed packet on the right. The electric field oscillation at optical frequencies is shown for illustrative purposes.

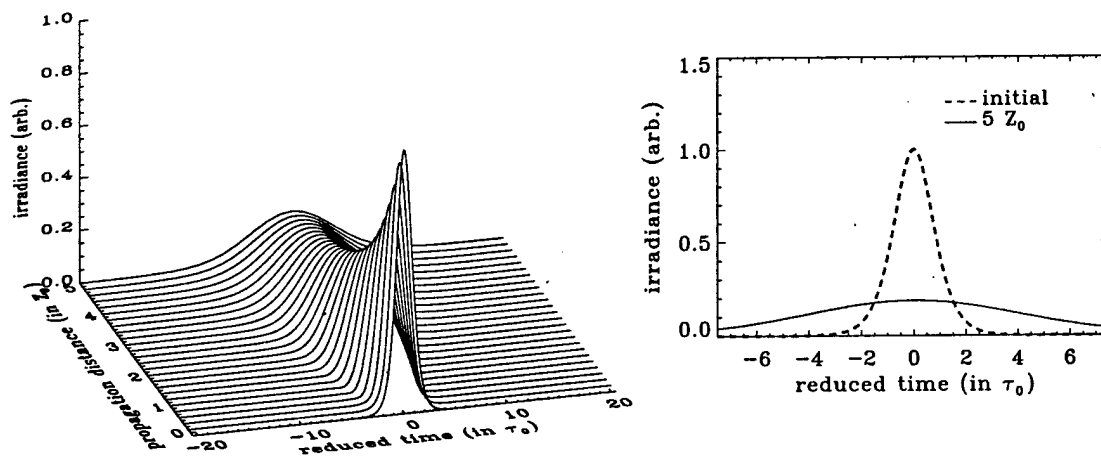


Figure 2.8: Dispersion of a pulse of initial sech(T/τ_0) amplitude profile, where $\tau_0 = 9.36$ fs and intensity FWHM duration 16.5 fs. After propagating $5 Z_0$, where $Z_0 = \pi\tau_0^2/2|k_0''| = 4.93$ mm $= 2L_D/\pi$ is the dispersive confocal distance and L_D is the standard dispersion length, the pulse FWHM broadens to 91.5 fs.

neglected. The dispersion of a wavepacket is illustrated in Figure 2.7 in the time domain. Even though group delay dispersion is an envelope phenomenon, the local phase of the envelope affects the underlying carrier. The figure indicates that in the dispersed envelope, higher frequencies move towards the pulse leading edge (negative time) and lower frequencies move towards the trailing edge (positive time), characteristic of the AGDD regime.

Figure 2.8 shows the effect of group-delay dispersion on a 16.5 fs FWHM pulse envelope over the dispersive equivalent distance of $5 Z_0$. Note that higher-order dispersive effects must be considered for this wavepacket, but are presently neglected for illustrative purposes. The definition of Z_0 is obtained by using the substitution $\tau_0 \rightarrow w_0\sqrt{k_0|k_0''|}$ in the definition of the confocal distance used for spatial diffraction of the sech(x/w_0) profile. With the consideration of only group-delay dispersion, the broadening of the temporal profile depends on the magnitude of GDD, not the sign. The chirp of the carrier does depend on the sign of GDD, as shown in Figure 2.7.

The nonlinear part of equation 2.52 has the solution

$$A(T, z) = e^{i\phi^{NL}(T, z)} A(0) \quad (2.55)$$

with the nonlinear phase accumulation

$$\phi^{NL}(T, z) = k_f n_2 \int_0^z |A(T, z')|^2 dz'. \quad (2.56)$$

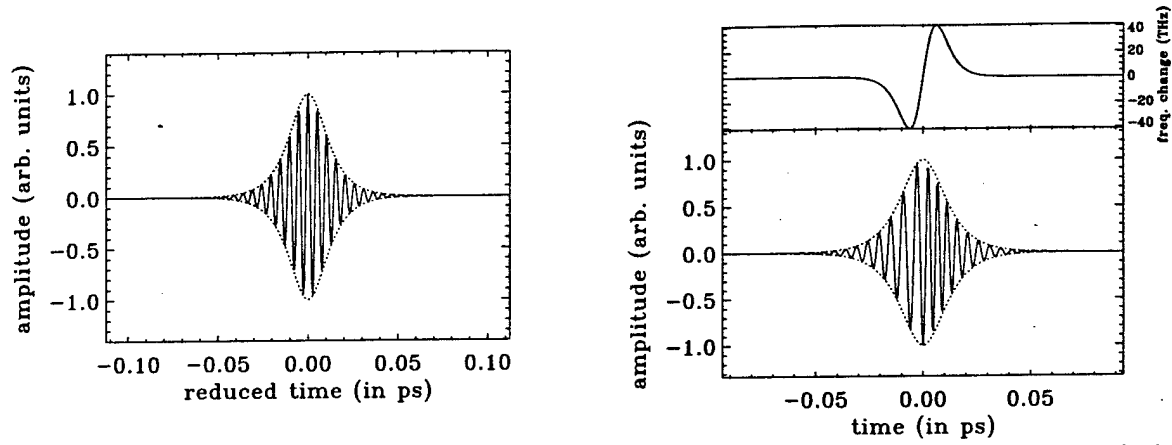


Figure 2.9: Effect of positive self-phase modulation on a 16.5 fs quasi-monochromatic wavepacket where the initial sech() packet is on the left and the nonlinearly modulated packet on the right, with π peak nonlinear phase change. Positive self-phase modulation reduces the phase velocity proportionally to the square of the field envelope resulting in an instantaneous frequency shift (top right), such that the pulse leading edge is downshifted and the pulse trailing edge is upshifted.

Because of the nonlinear phase modulation, there is an instantaneous frequency change across the pulse envelope. This frequency shift is defined as

$$\delta\omega = -\frac{\partial\phi^{\text{NL}}}{\partial T} \approx -k_f n_2 \Delta z \frac{\partial |A(T, \Delta z/2)|^2}{\partial T}, \quad (2.57)$$

and is known as self-phase modulation (SPM) [30]. This effect is illustrated in Figure 2.9. The frequencies at the center of the pulse are not shifted because the derivative there is zero. At the inflection points of the envelope the derivative is maximum resulting in the largest frequency shift such that the leading edge is downshifted while the trailing edge is upshifted for positive SPM. Therefore, in the anomalous dispersion regime, the leading edge will travel with a slower group velocity than the trailing edge. The net effect is a narrowing, or compression, of the pulse and a temporal soliton results from the dynamical balance between positive SPM and linear dispersion in the anomalous regime. This balance is the reason for the existence of stable temporal solitons. In the normal regime, pulse lengthening is enhanced due to positive SPM.

It is clear from Figure 2.9 that, even though the envelope amplitude is unchanged, SPM alters the phase and hence the spectral content. This phenomenon is known as SPM-induced spectral broadening and is shown in Figure 2.10 by comparing the temporal Fourier transforms of the two wavepackets in Figure 2.9, where significant spectral broadening has occurred. In the absence of asymmetric, higher-order effects, spectral broadening is symmetric about the central frequency.

By analogy with the 1-D spatial soliton, the existence of a temporal soliton of the form

$$A(T, z) = A_0 \operatorname{sech}\left(\frac{T}{\tau_0}\right) e^{i\beta z} \quad (2.58)$$

is postulated where β is the nonlinear wave number correction. Using this trial solution in equation 2.52 results in

$$A_0^2 = -\frac{k_0'' n_0}{k_0 n_2 \tau_0^2} \quad (2.59)$$

$$\beta = -k_0''/2\tau_0^2 = k_f n_2 A_0^2/2. \quad (2.60)$$

In order for the amplitude A_0 to be real, $k_0'' < 0$, which is the anomalous dispersion regime as anticipated. By comparing the expression for the amplitude in the temporal and spatial cases, there is a correspondence between the temporal duration and spatial width give by $\tau_0 = \sqrt{-k_0 k_0''} w_0$, also indicating that $k_0'' < 0$ in order for τ_0 to be real.

The full SVEA temporal soliton solution is then

$$\bar{A}(x, y, z, T) = \frac{1}{\tau_0} \sqrt{\frac{-k_0'' n_0}{k_0 n_2}} \Phi(x, y) \operatorname{sech}\left(\frac{T}{\tau_0}\right) e^{i[k_0 - k_0''/2\tau_0^2]z}, \quad (2.61)$$

where the guided profile $\Phi(x, y)$ and linear wave number k_0 are added for completeness. The numerical propagation of a temporal soliton is shown in Figure 2.11 over 5 Z_0 . As for the spatial soliton, the pulse duration and position in the reduced time coordinate frame are unchanged.

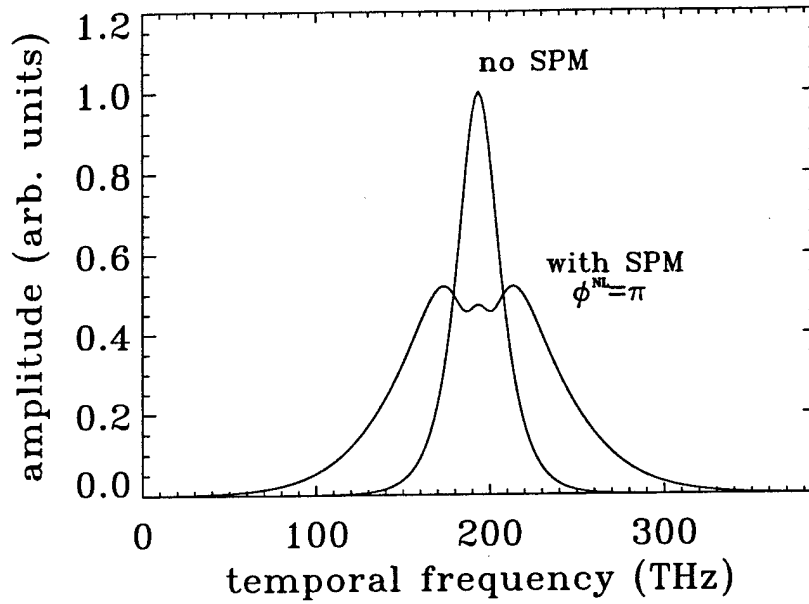


Figure 2.10: Self-phase modulation induced spectral broadening. The peak nonlinear phase change is π , resulting in spectral broadening by greater than a factor of 2. Subsequent pulse compression utilizing the appropriate net negative group-delay dispersion will result in a pulse shorter than the initial pulse.

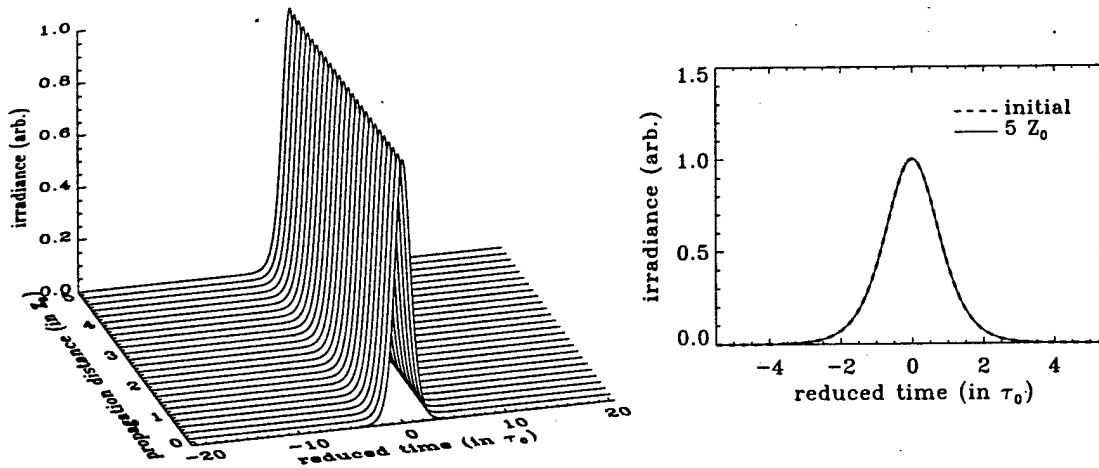


Figure 2.11: Temporal soliton pulse of initial $\text{sech}(T/\tau_0)$ profile, where $\tau_0 = 9.36$ fs and intensity FWHM duration 16.5 fs. After propagating $5 Z_0$, where $Z_0 = 4.93$ mm, the pulse duration and position are unchanged.

The optical intensity is given by

$$I(x, y, z, T) = -\frac{\epsilon_0 c n(x, y) k_0''}{2k_f \tau_0^2 n_2} \Phi^2(x, y) \text{sech}^2\left(\frac{T}{\tau_0}\right), \quad (2.62)$$

where the spatial variation of the linear refractive index responsible for transverse guidance is retained. The optical power of the soliton is

$$P(z, T) = -\frac{\epsilon_0 c n_0 k_0'' A_{\text{eff}}}{2k_f \tau_0^2 n_2} \text{sech}^2\left(\frac{T}{\tau_0}\right), \quad (2.63)$$

and the effective transverse spatial area is defined by the expression

$$A_{\text{eff}} = \frac{1}{n_0} \int \int n(x, y) \Phi^2(x, y) dx dy. \quad (2.64)$$

The total energy of the soliton is then

$$E = -\frac{\epsilon_0 c n_0 k_0'' A_{\text{eff}}}{k_f \tau_0 n_2}, \quad (2.65)$$

with the same scaling as the 1-D spatial soliton.

As was done for the spatial paraxial soliton solution, the temporal SVEA soliton solution can be generalized to the form

$$A(z, T) = A_0 \operatorname{sech} \left(\frac{T - \delta k_0' z}{\tau_0} \right) e^{i[\beta z - \delta \omega T]}. \quad (2.66)$$

Substituting into the temporal NLS equation 2.52 results in

$$A_0^2 = -\frac{k_0'' n_0}{k_0 n_2 \tau_0^2} \quad (2.67)$$

$$\beta = -\frac{k_0''}{2\tau_0^2} + \frac{\delta \omega^2 k_0''}{2} \quad (2.68)$$

$$\delta k_0' = \delta \omega k_0''. \quad (2.69)$$

This result simply means that a change in optical frequency $\delta \omega$ changes the linear wave number from k_0 to $k_0 + \delta \omega^2 k_0''/2$ and the group delay from k_0' to $k_0' + \delta \omega k_0''$. Notice that the amplitude is unchanged which is the result of neglecting the third-order dispersion contribution k_0''' in the NLS equation.

2.4.3 Higher-Order 1-D Temporal NLS Solitons

Taking the temporal nonlinear wave equation one order of approximation beyond SVEA results in the modified NLS, or mNLS, equation

$$2ik_0 \frac{\partial A}{\partial z} - k_0 k_0'' \frac{\partial^2 A}{\partial T^2} - \frac{i}{3} k_0 k_0''' \frac{\partial^3 A}{\partial T^3} + 2k_0^2 \frac{n_2}{n_0} |A|^2 A + 4ik_f n_2 \left[\frac{n_0}{c} - \frac{k_0'}{2} \right] \frac{\partial |A|^2 A}{\partial T} = 0. \quad (2.70)$$

A similar equation was previously considered [140] but lacked the correct form of the shock term. As shown in Chapter 3, this is the order at which the effects of stimulated Raman scattering also appear for the given scalings. Raman scattering is a dissipative effect and is considered in the next section on higher-order temporal effects.

A solution to the mNLS equation is given by

$$A(z, T) = A_0 \operatorname{sech} \left(\frac{T - \delta k_0' z}{\tau_0} \right) e^{i[\beta z - \delta \omega T]}, \quad (2.71)$$

with the following parameter values

$$A_0^2 = -\frac{k_0 k_0'''}{6k_f n_2 \tau_0^2 [n_0/c - k_0'/2]} \quad (2.72)$$

$$\beta = -\frac{1}{2\tau_0^2} [k_0'' + \delta \omega k_0'''] + \frac{\delta \omega^2 k_0''}{2} + \frac{\delta \omega^3 k_0'''}{6} \quad (2.73)$$

$$\delta k_0' = \delta \omega k_0'' + \frac{\delta \omega^2 k_0'''}{2} - \frac{k_0'''}{6\tau_0^2} \quad (2.74)$$

$$\delta \omega = \frac{k_0}{4[n_0/c - k_0'/2]} - \frac{3k_0''}{2k_0'''} \quad (2.75)$$

The only free parameter is the pulse duration τ_0 , which allows for multiple solitons of different durations/amplitudes to be used for switching. Stability has not been investigated theoretically, but numerical simulation verifies that the solution propagates

stably. Note that this solution is a generalization of that obtained previously [140] in which temporal walkoff in the reduced time coordinate frame was not allowed, resulting in no free parameters. This is explained further below.

There are two special cases of interest in which $\delta k'_0 = 0$ and $\delta\omega = 0$. When $\delta k'_0 = 0$, there is no pulse walkoff in the reduced-time coordinate frame. In this case there are no free parameters [140] since the pulse duration is now fixed at the value

$$\tau_0|_{\delta k'_0=0} = \frac{k_0'''}{6[\delta\omega k_0'' + \frac{1}{2}\delta\omega^2 k_0''']}. \quad (2.76)$$

Another way of expressing this relationship is to determine the value of TOD needed to sustain a given pulse duration. Therefore, choosing the pulse duration τ_0 as a "free" parameter, the following conditions hold

$$\delta\omega|_{\delta k'_0=0} = \frac{2\tau_0 k_0 - \sqrt{4k_0^2 \tau_0^2 - 16[n_0/c - k'_0/2]^2}}{4\tau_0 [n_0/c - k'_0/2]} \quad (2.77)$$

$$k_0'''|_{\delta k'_0=0} = -\frac{6\delta\omega k_0'' \tau_0^2}{3\delta\omega^2 \tau_0^2 - 1}, \quad (2.78)$$

where the '-' sign was chosen in front of the square-root as the proper solution which is verified by numerical simulation.

The other interesting case is when $\delta\omega = 0$, meaning that the original choice of center frequency ω_0 is a solution to the mNLS equation. This can only occur when the material dispersion parameters satisfy the condition

$$k_0'''|_{\delta\omega=0} = \frac{6k_0'' [n_0/c - k'_0/2]}{k_0}, \quad (2.79)$$

such that the change in group delay is written

$$\delta k'_0|_{\delta\omega=0} = -\frac{k_0'' [n_0/c - k'_0/2]}{k_0 \tau_0^2}. \quad (2.80)$$

Note that τ_0 remains a free parameter. The temporal walkoff in the reduced time coordinate frame can be eliminated by redefining the coordinate transformation as $T = t - (k'_0 + \delta k'_0)z$.

A special case of mNLS, written in normalized form, is known to be integrable [208]

$$i\frac{\partial u}{\partial z} + \frac{1}{2}\frac{\partial^2 u}{\partial T^2} + |u|^2 u + i\beta \left[\frac{\partial^3 u}{\partial T^3} + 6|u|^2 \frac{\partial u}{\partial T} \right] = 0 \quad (2.81)$$

because it is a member of the NLS hierarchy and has an associated scattering problem and inverse scattering transform. This equation is not of interest here because, like the general mNLS solution with $\delta\omega = 0$, it places restrictions on the material constants which may not be feasible to satisfy. Instead, this thesis deals with more general, non-integrable cases using parameters of materials currently in use for nonlinear optical switching.

2.4.4 Higher-Order Temporal Effects

The previous section considered additional non-SVEA terms to the NLS equation that still allowed for analytic soliton-like solutions. This section considers additional terms that result from the scalar, temporal reduction of the full evolution equation derived in Chapter 3 for use in spatio-temporal propagation. The additional terms included here that are not present in equation 2.70 are fourth-order dispersion (FOD) and stimulated Raman scattering:

$$\begin{aligned} 2ik_0 \frac{\partial A}{\partial z} - k_0 k_0'' \frac{\partial^2 A}{\partial T^2} - \frac{i}{3} k_0 k_0''' \frac{\partial^3 A}{\partial T^3} + \frac{1}{12} k_0 k_0'''' \frac{\partial^4 A}{\partial T^4} \\ + 2k_0^2 \frac{n_K}{n_0} |A|^2 A + 4ik_f n_K \left[\frac{n_0}{c} - \frac{k'_0}{2} \right] \frac{\partial |A|^2 A}{\partial T} \\ + \frac{k_f}{2} \left\{ k_f + i \frac{2}{n_0} \left[\frac{n_0}{c} - \frac{k'_0}{2} \right] \frac{\partial}{\partial T} \right\} \int_0^\infty R_R(\tau) |A(T-\tau)|^2 A d\tau = 0. \end{aligned} \quad (2.82)$$

Here, k_0'''' is the fourth-order dispersion coefficient, n_K is the nonlinear Kerr index, and $R_R(\tau)$ is the Raman response function. To date, there is no known analytic solution to this equation. As a result, the effects of these higher-order temporal terms on the fundamental NLS soliton are now discussed.

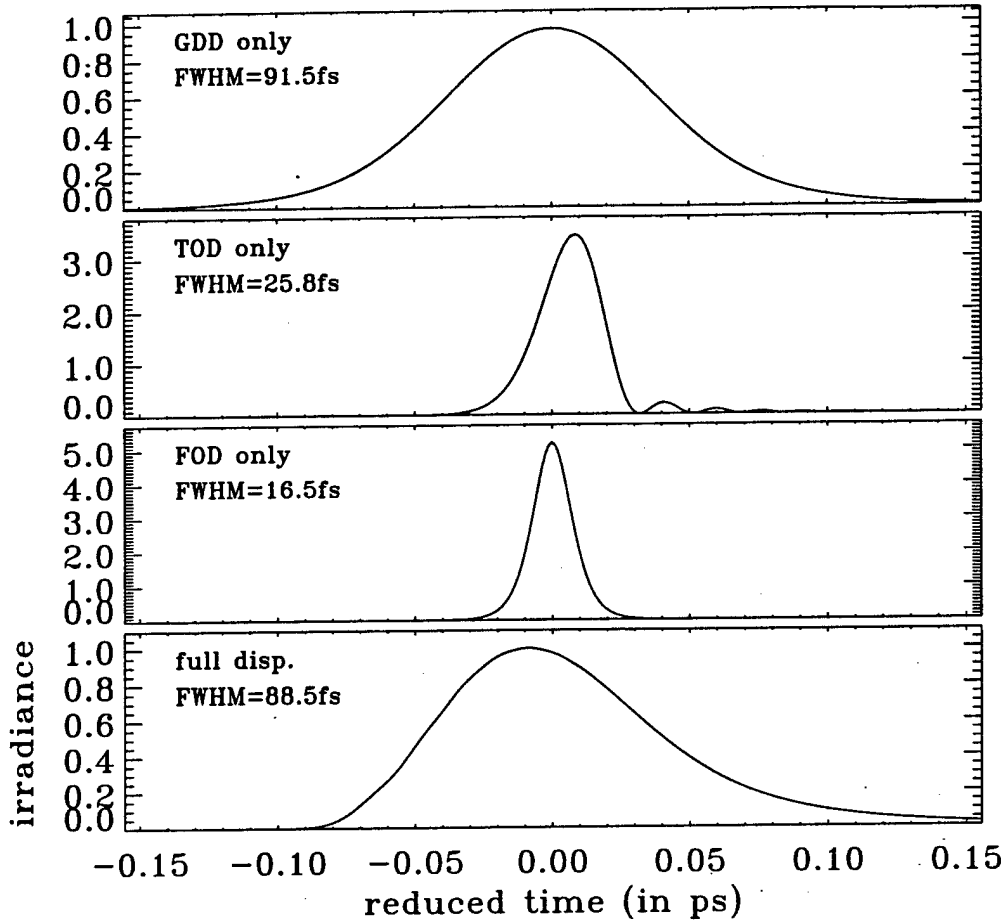


Figure 2.12: Effects of linear dispersion on a sech() pulse with initial intensity temporal FWHM of 16.5 fs. Based on the dispersion properties of fused silica, the group-delay dispersion (GDD) coefficient $k_0'' < 0$, the third-order dispersion (TOD) coefficient $k_0''' > 0$, and the fourth-order dispersion (FOD) coefficient $k_0'''' < 0$. The propagation distance in each figure is $5 Z_0$.

Higher-Order Linear Dispersion

First, the effects of higher-order linear dispersion will be considered in the absence of nonlinearity. In this case, equation 2.82 reduces to

$$2ik_0 \frac{\partial A}{\partial z} - k_0 k_0'' \frac{\partial^2 A}{\partial T^2} - \frac{i}{3} k_0 k_0''' \frac{\partial^3 A}{\partial T^3} + \frac{1}{12} k_0 k_0'''' \frac{\partial^4 A}{\partial T^4} = 0. \quad (2.83)$$

The group delay as a function of frequency, as shown in Figure 2.6, can be written $k'(\omega) = k'_0 + \Delta k'(\Delta\omega)$, where the approximation

$$\Delta k'(\Delta\omega) \approx \Delta\omega k_0'' + \frac{1}{2} \Delta\omega^2 k_0''' + \frac{1}{6} \Delta\omega^3 k_0''', \quad (2.84)$$

is made and is cubic in frequency deviation about ω_0 .

The effects of linear dispersion on a temporal wavepacket are shown in Figure 2.12. The top plot shows temporal broadening due to GDD alone, and is based on the same data used for Figure 2.8. The symmetry of the broadening can be understood by considering the change in group delay, which is linear in $\Delta\omega$. Therefore, symmetric deviation about the center frequency results in anti-symmetric deviation in group delay about k'_0 , such that

$$k'(\omega_0 \pm \Delta\omega) \approx k'_0 \pm \Delta\omega k_0''. \quad (2.85)$$

As a result, GDD leads to symmetric broadening, with red-shifted frequencies trailing blue-shifted frequencies in the AGDD regime.

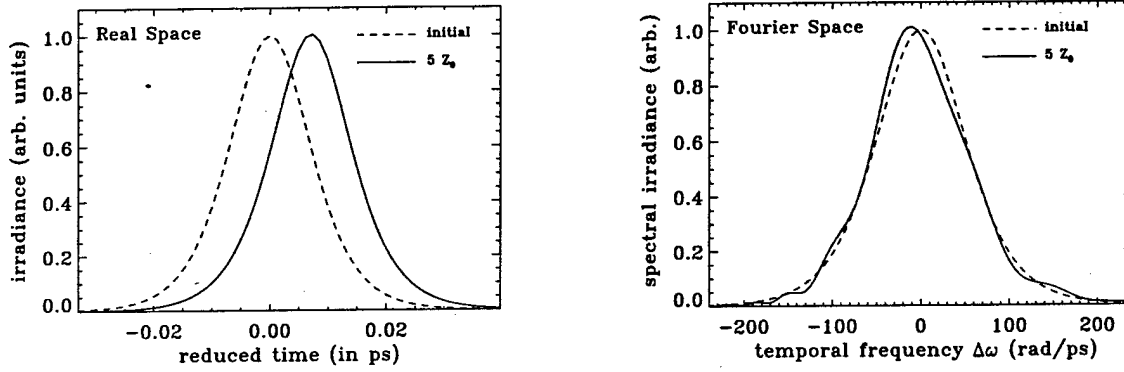


Figure 2.13: Effects of higher-order linear dispersion on a temporal soliton with initial intensity temporal FWHM of 16.5 fs. The initial (dashed curve) and final (solid curve, after $5 Z_0$) temporal envelopes are shown on the left plot, while the initial (dashed) and final (solid) spectral envelopes are shown on the right. The soliton is delayed by 7.25 fs and narrows slightly to 16.2 fs.

The second plot shows temporal broadening due to TOD alone. In this case, the group delay

$$k'(\omega_0 \pm \Delta\omega) \approx k'_0 + \frac{1}{2} \Delta\omega^2 k_0''', \quad (2.86)$$

which results in a symmetric change in group delay with frequency. This expression indicates that both red-shifted and blue-shifted frequencies will move towards the leading edge ($\text{TOD} < 0$) or trailing edge ($\text{TOD} > 0$). For positive TOD as shown, the peak will be delayed in the reduced time coordinate frame with steepening of the trailing edge and oscillations in the tail [30].

The third plot shows the effect of FOD alone. The group delay can be written

$$k'(\omega_0 \pm \Delta\omega) \approx k'_0 \pm \frac{1}{6} \Delta\omega^3 k_0''', \quad (2.87)$$

which, like GDD, is anti-symmetric. Again, broadening will be symmetric, but with much stronger change in group delay for frequencies far removed from ω_0 . FOD can either enhance (same sign) or oppose (opposite sign) the broadening due to GDD. As shown in the plot, the FOD coefficient of fused silica is small enough that its effect is negligible for the choice of initial pulse duration and propagation distance.

Finally, the fourth plot shows the temporal broadening due to the combined effects of all dispersion terms. In fact, the full dispersion relation using the Sellmeier coefficients for fused silica (see Appendix D) is used, but the primary effects arise from GDD and TOD. Because of TOD, broadening is asymmetric. The signs of GDD (< 0) and TOD (> 0) indicate that TOD opposes the change in group delay due to GDD when $\Delta\omega > 0$, and enhances the change due to GDD when $\Delta\omega < 0$, as shown in Figure 2.6. Therefore, the broadening of the red-shifted trailing edge will be increased while broadening of the blue-shifted leading edge will be reduced. Opposite to the case of TOD alone, the peak is now slightly advanced.

The effect of higher-order dispersion on a temporal soliton is shown in Figure 2.13, as described by the evolution equation

$$2ik_0 \frac{\partial A}{\partial z} - k_0 k_0'' \frac{\partial^2 A}{\partial T^2} - \frac{i}{3} k_0 k_0''' \frac{\partial^3 A}{\partial T^3} + \frac{1}{12} k_0 k_0'''' \frac{\partial^4 A}{\partial T^4} + 2k_0^2 \frac{n_2}{n_0} |A|^2 A = 0, \quad (2.88)$$

where $n_2 = n_K + \frac{1}{4n_0} \int_0^\infty R_R(\tau) d\tau$ is the total instantaneous nonlinear refractive index. Because AGDD is balanced with self-phase modulation, the soliton is delayed by 7.25 fs in the reduced coordinates. This delay is primarily the result of positive TOD, and is similar to the case previously discussed with TOD acting alone on a linear pulse, where the delay was 7.08 fs. The peak of the spectrum is also downshifted, as shown on the right-hand plot, which corresponds to the delay experienced in time. Except for the delay, in real space, the soliton is mostly unaffected by higher-order dispersion, a testament to (1+1)-D stability.

Optical Shock

Optical shock [183] is the first-order time derivative of the nonlinearity and gives rise to an intensity-dependent group delay [184]. The evolution equation describing nonlinear refraction and optical shock is

$$2ik_0 \frac{\partial A}{\partial z} + 2k_0^2 \frac{n_2}{n_0} |A|^2 A + 4ik_f n_2 \left[\frac{n_0}{c} - \frac{k'_0}{2} \right] \frac{\partial |A|^2 A}{\partial T} = 0. \quad (2.89)$$

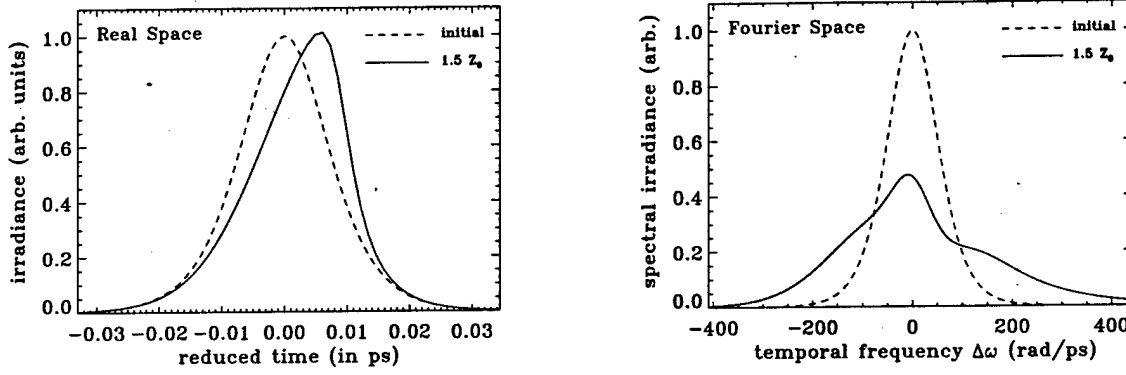


Figure 2.14: Effects of optical shock on a 16.5 fs wavepacket for a propagation distance of $1.5 Z_0$. The left-hand plot shows the initial (dashed) and final (solid) temporal intensity envelopes and indicates slight narrowing to 16.3 fs and delay of the peak by 5.68 fs. The right-hand plot shows the initial (dashed) and final (solid) spectral envelopes and indicates spectral broadening from 19.1 THz to 33.9 THz due to SPM with asymmetry caused by shock.

The optical shock expression obtained here has two contributions, one (proportional to phase delay) directly from Maxwell's equation in the time derivative of the nonlinear polarization, and one (proportional to group delay) due to the removal of the SVE approximation by one order.

Expressing the field envelope as

$$A = |A|e^{i\phi}, \quad (2.90)$$

equation 2.89 can be written as the coupled differential equations

$$\frac{\partial |A|}{\partial z} = 6 \frac{n_2}{n_0} \left[\frac{k'_0}{2} - \frac{n_0}{c} \right] |A|^2 \frac{\partial |A|}{\partial T} \quad (2.91a)$$

$$\frac{\partial \phi}{\partial z} = k_f n_2 |A|^2 + \frac{n_2}{n_0} \left[\frac{k'_0}{2} - \frac{n_0}{c} \right] |A|^2 \frac{\partial \phi}{\partial T}. \quad (2.91b)$$

The first equation has the simple solution

$$|A(T, z)| = |A(z - v_{\text{igv}} T)|, \quad (2.92)$$

where the intensity-dependent group velocity is defined

$$v_{\text{igv}} = \frac{1}{6 \frac{n_2}{n_0} \left[\frac{k'_0}{2} - \frac{n_0}{c} \right] |A|^2}. \quad (2.93)$$

Therefore, the group velocity is reduced with increasing intensity.

The self-steepening effect due to optical shock is shown in Figure 2.14. The peak is delayed more than the wings resulting in a steepening of the trailing edge. An estimate of the time delay of the peak is

$$\Delta T = \frac{dZ_0}{v_{\text{igv}}} = \frac{3}{2} \lambda d \left[\frac{n_0}{c} - \frac{k'_0}{2} \right], \quad (2.94)$$

which gives the value $\Delta T = 5.73$ fs using the parameters of the simulations. This compares very well with the value of 5.68 fs obtained by numerically locating the peak. The second plot shows that the spectral broadening due to self-phase modulation (as shown in Figure 2.10) becomes asymmetric, with slight down-shift of the peak corresponding to delay in the temporal domain.

The steepening of the trailing edge results in broad frequency content. Therefore, dispersion must be considered, which dissipates steepening. The following evolution equation well describes the effects of shock and dispersion

$$\begin{aligned} 2ik_0 \frac{\partial A}{\partial z} - k_0 k_0'' \frac{\partial^2 A}{\partial T^2} - \frac{i}{3} k_0 k_0''' \frac{\partial^3 A}{\partial T^3} + \frac{1}{12} k_0 k_0'''' \frac{\partial^4 A}{\partial T^4} \\ + 2k_0^2 \frac{n_K}{n_0} |A|^2 A + 4ik_f n_K \left[\frac{n_0}{c} - \frac{k'_0}{2} \right] \frac{\partial |A|^2 A}{\partial T} = 0, \end{aligned} \quad (2.95)$$

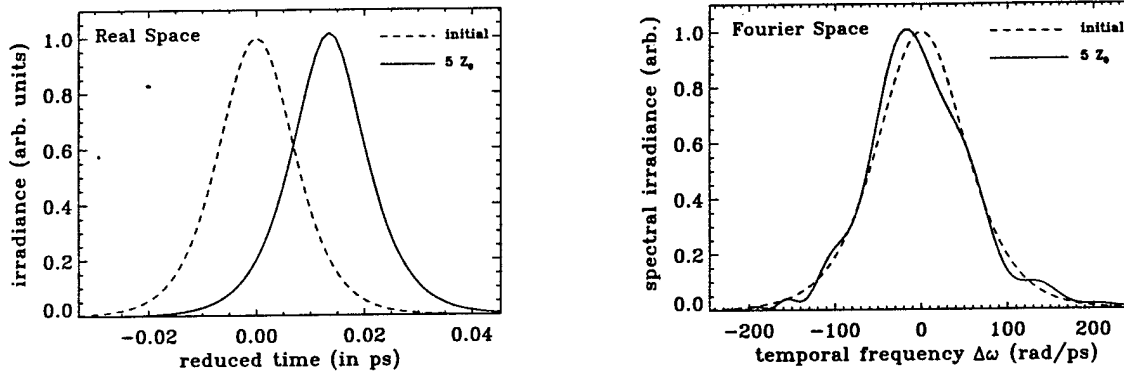


Figure 2.15: Effects of optical shock and higher-order dispersion on a 16.5 fs wavepacket for a propagation distance of $5 Z_0$. The left-hand plot shows the initial (dashed) and final (solid) temporal intensity envelopes and indicates slight narrowing to 15.8 fs and overall (centroid) delay of 13.7 fs. The right-hand plot shows the initial (dashed) and final (solid) spectral envelopes and indicates asymmetry due to shock and TOD.

and is used for the simulation results presented in Figure 2.15, using the fundamental soliton $A = \text{sech}(T/\tau_0)$ as the initial condition. Dispersion dissipates the shock and the balance between AGDD and SPM stabilizes the pulse. As a result, the effect of shock is to delay the entire pulse. The overall delay of 15.8 fs is due to both shock and positive TOD, where the delay without shock is 7.25 fs, as shown in Figure 2.13.

Raman Scattering

This section considers the full evolution equation 2.82, which includes higher-order dispersion, shock and Raman scattering. The Raman nonlinearity is described via the full Raman response function [186] for a single resonance:

$$R_R(\tau) = \frac{R_0 e^{-\gamma\tau/2} \sin(\Omega_R \tau)}{\Omega_R}, \quad (2.96)$$

where $\Omega_R = \sqrt{\Omega_f^2 - \gamma^2/4}$ is the optical phonon frequency, Ω_f is the natural oscillation frequency and γ is the damping constant. Raman scattering results in a continuous frequency down-shift by providing gain at lower frequencies at the expense of higher frequencies, as discussed in Appendix B. This effect is often referred to as the soliton self-frequency shift [145]. In the AGDD regime, the downshifted frequencies travel with a greater group delay (slower group velocity) thereby delaying the wave in the reduced time coordinates, as shown in Figure 2.16.

Spectral narrowing and downshift are clearly seen on the right-hand plot. The peak of the Raman gain in fused silica is 13.2 THz (or 82.9 rad/ps). The peak of the downshifted spectrum is about 80 rad/ps below the knee of the high frequency side of the curve, indicating that the low frequencies are amplified and the high frequencies are attenuated, resulting in spectral narrowing. Spectral four-wave mixing continuously creates new frequencies, called the Stokes (low-frequency) and anti-Stokes (high-frequency) side bands, allowing for a Raman downshift which exceeds the original spectral bandwidth. The manifestation of these effects in the temporal domain is temporal broadening (due to spectral narrowing) and delay (due to downshift). Note that the delay of 20.0 fs is only slightly greater than the delay of 13.7 fs without the Raman term. This indicates that, in this case, the Raman effect is about the same order as the other higher-order temporal effects, but because of the large, continuous spectral downshift, the Raman term will manifest itself over much longer distances and eventually dominate.

2.5 2-D and 3-D Spatio-Temporal Solitary Waves

The propagation and interaction of multi-dimensional spatio-temporal solitary waves, which are stationary in both space and time, is of ultimate interest to this thesis. These multi-dimensional nonlinear phenomena are termed solitary waves instead of solitons because, like solitons, they are stationary, but may not satisfy the additional properties of solitons such as inelastic collisions (which preserve the soliton eigenvalues) and integrability of the defining equations. Like temporal solitons in fiber, these multi-dimensional solitary waves are fully confined, either by complete nonlinear self-confinement in the case of the 3-D light bullet or with one dimension of linear confinement by a planar waveguide in the 2-D self-confined case, such that the

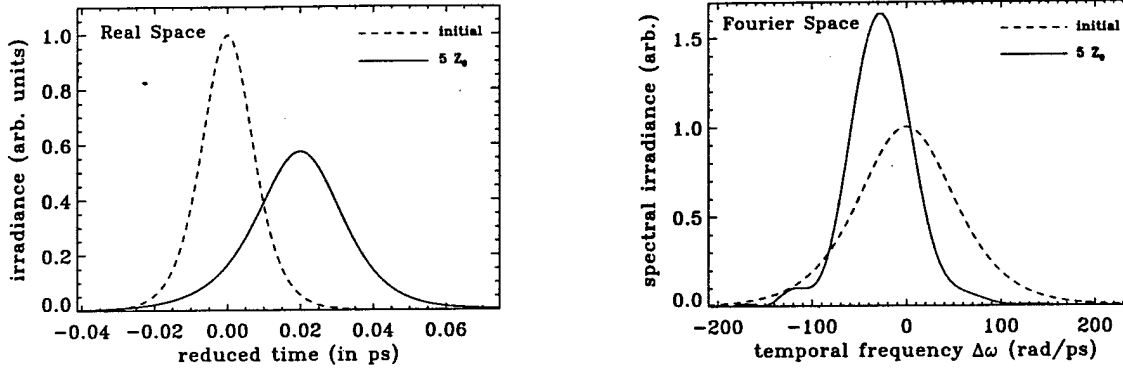


Figure 2.16: Effects of Raman scattering, including shock and higher-order dispersion, on a 16.5 fs wavepacket for a propagation distance of $5 Z_0$. The left-hand plot shows the initial (dashed) and final (solid) temporal intensity envelopes and indicates broadening to 27.4 fs and overall (centroid) delay of 20.0 fs. The right-hand plot shows the initial (dashed) and final (solid) spectral envelopes and indicates spectral narrowing from 19.1 THz to 11.7 THz and overall downshift by 4.23 THz.

energy requirement is low. A brief discussion of 2-D spatio-temporal solitary waves is presented here with more detail on propagation and interaction provided in Chapter 6.

The most straightforward means of describing the propagation of 2-D spatio-temporal solitary waves is the direct extension of the (1+1)-D nonlinear Schrödinger (NLS) equation

$$2ik_0 \frac{\partial A}{\partial z} + \frac{\partial^2 A}{\partial x^2} - k_0 k_0'' \frac{\partial^2 A}{\partial T^2} + 2k_0^2 \frac{n_2}{n_0} |A|^2 A = 0, \quad (2.97)$$

but it is well known that the solitary wave solutions are unstable to propagation [209]. This instability can result in critical collapse, or blow-up, when nonlinearity overtakes diffraction/dispersion, or in broadening when diffraction/dispersion overtakes nonlinearity. In the case of (2+1)-D spatial and spatio-temporal in the AGDD regime, and (3+1)-D spatio-temporal in the AGDD regime, blow-up is the result [176], which indicates that the evolution equation does not take into account additional physics which arrest this behavior. The inclusion of non-paraxial and non-SVEA terms results instead in spatio-temporal broadening.

Since the pump must propagate the length of the logic gate undisturbed (i.e. with final size about the same as the initial size), mechanisms to stabilize propagation need to be investigated. Multi-dimensional propagation can be stabilized by the inclusion of a suitable nonlinear index saturation mechanism [2, 210] such as ultrafast quintic nonlinearity of opposite sign to the cubic Kerr nonlinearity, i.e. $n = n_0 + n_2 |A|^2 + n_4^{\text{eff}} |A|^4$ where A is a scalar field and $n_4^{\text{eff}} < 0$ accounts for the refractive part of the effective quintic nonlinearity. Examples of this type of saturation behavior are the nonresonant nonlinearities of semiconducting AlGaAs [143] at sub-half bandgap or the organic single-crystal PTS [125, 147]. Ultrafast saturation from the quintic nonlinearity is the stabilization mechanism discussed in this thesis, but it should be noted that the balancing between third-order dispersion (plus space-time focusing) and shock as shown in section 2.4.3 for 1-D temporal solitons may also serve as a stabilization mechanism for multi-dimensional propagation [139].

The scalar (2+1)-D cubic-quintic NLS equation is

$$2ik_0 \frac{\partial A}{\partial z} + \frac{\partial^2 A}{\partial x^2} - k_0 k_0'' \frac{\partial^2 A}{\partial T^2} + 2k_0^2 \frac{n_2}{n_0} \left[|A|^2 + \frac{n_4^{\text{eff}}}{n_2} |A|^4 \right] A = 0. \quad (2.98)$$

Along with the quintic term, any additional terms to the (2+1)-D nonlinear evolution equation that have comparable effect must also be considered, and the conditions determined under which they can be neglected. This motivates the multiple-scales derivation of Chapter 3, in which a full vectorial derivation directly from Maxwell's equation is performed and results in a coupled, non-paraxial, non-slowly-varying envelope (SVE), first-order vectorial differential equation for the propagation of orthogonal linear polarizations including nonlinear couplings with the weak longitudinally-projected field. In Chapter 6, reduced, paraxial, versions of these equations are used to study the propagation of a single pump solitary wave and the vectorial interaction between pump and signal solitary waves. It is shown that quintic index saturation can stabilize against the effects of other, comparable, higher-order terms.

The scalar equation suitable for the propagation of a single spatio-temporal nonlinear wave under the conditions of present

interest is, from section 3.2.3,

$$\begin{aligned}
 & 2ik_0 \frac{\partial A}{\partial z} + \frac{\partial^2 A}{\partial x^2} - k_0 k_0'' \frac{\partial^2 A}{\partial T^2} - \frac{i}{3} k_0 k_0''' \frac{\partial^3 A}{\partial T^3} - i \frac{k_0'}{k_0} \frac{\partial^3 A}{\partial T \partial x^2} \\
 & + \frac{1}{12} k_0 k_0'''' \frac{\partial^4 A}{\partial T^4} + 2k_0^2 \frac{n_K}{n_0} |A|^2 A + 4ik_f n_K \left[\frac{n_0}{c} - \frac{k_0'}{2} \right] \frac{\partial |A|^2 A}{\partial T} \\
 & + \frac{k_f}{2} \left\{ k_f + i \frac{2}{n_0} \left[\frac{n_0}{c} - \frac{k_0'}{2} \right] \frac{\partial}{\partial T} \right\} \int_0^\infty R_R(\tau) |A(T-\tau)|^2 A d\tau \\
 & + 2k_0^2 \frac{n_4^{\text{eff}}}{n_0} |A|^4 A = 0.
 \end{aligned} \tag{2.99}$$

The first term in equation 2.99 is the first-order propagator along the z direction. The reduced time transformation $T = t - k_0' z$, where k_0' is the group-delay, is made such that propagation is nominally stationary in the new coordinate system. Paraxial diffraction along the x direction and group-delay dispersion are represented by the next two terms, where k_0'' is the group-delay dispersion coefficient and describes pulse lengthening under the SVE approximation. Third-order dispersion, with coefficient k_0''' , and space-time focusing [144] follow, and are the first non-SVE corrections. The space-time focusing term describes the (paraxial) curvature of the energy front due to spatio-temporal diffraction. The final linear term is fourth-order dispersion.

The first nonlinear term in equation 2.99 represents third-order nonlinear refraction followed by optical shock. The following terms describe the Raman nonlinearity. The effective quintic nonlinear index n_4^{eff} consists of three distinct contributions: directly from the fifth-order polarization, from the product of the third-order polarization with itself due to the reduction from Maxwell's equations to a first-order equation, and a "cascaded" contribution due to the nonlinear coupling between the fundamental and third-harmonic, which can be tuned via phase-matching [211]. A simple estimate of the size of the cascaded quintic nonlinear index suggests that [134] $n_4^{\text{eff}} \approx -20n_2^2$, which is negative as desired for multi-dimensional stability.

The initial conditions used in the simulations of Chapter 6 are the numerically-computed eigenmodes of the normalized scalar (2+1)-D cubic-quintic NLS equation:

$$i \frac{\partial u}{\partial \tilde{z}} + \frac{\partial^2 u}{\partial \tilde{x}^2} + s \frac{\partial^2 u}{\partial \tilde{T}^2} + 2|u|^2 u + 2q|u|^4 u = 0, \tag{2.100}$$

where the following definitions are made: $u = k_0 w_0 \sqrt{n_2/n_0} A$ where $n_2 = n_K + \frac{1}{4n_0} \int_0^\infty R_R(\tau) d\tau$ is the total instantaneous nonlinear refractive index, $\tilde{z} = z/2k_0 w_0^2$, $\tilde{x} = x/w_0$, $\tilde{T} = T/w_0 \sqrt{|k_0 k_0''|}$, $s = -\text{sign}(k_0'')$, and $q = n_0 n_4^{\text{eff}} / n_2^2 k_0^2 w_0^2$ with w_0 a measure of the transverse spatial width of the solitary-wave. For a bright spatio-temporal solitary wave with $n_2 > 0$, operation must be in a region of anomalous group-delay dispersion (AGDD) such that $s = +1$. The symmetry of this cubic-quintic equation allows it to be transformed into an ordinary differential equation [2] and subsequently solved via standard fourth-order Runge-Kutta [212].

The existence of a radially-symmetric stationary solution to equation 2.100 of the form

$$u(\tilde{y}, \tilde{T}, \tilde{z}) = U(\rho) \exp(i\beta \tilde{z}) \tag{2.101}$$

is postulated, where $\rho = \sqrt{\tilde{y}^2 + \tilde{T}^2}$. Substituting this ansatz into equation 2.100 results in the ordinary nonlinear differential equation

$$\frac{d^2 U}{d\rho^2} + \frac{1}{\rho} \frac{dU}{d\rho} + [2U^2 + 2qU^4 - \beta] U = 0, \tag{2.102}$$

which has fundamental and higher-order solutions corresponding to increasing optical energies, physical size, and number of zero crossings of the field [2].

For the purposes of optical switching, only the fundamental eigenmode solutions to equation 2.102 are of interest, which possess the minimum size and energy [2]. It should also be noted that the higher-order modes are unstable to angular perturbations even with a saturating nonlinearity [213]. There is a family of such fundamental solutions of different widths and amplitudes parameterized by the value of q . Using the relationship between the field amplitude A and the normalized amplitude u , the q parameter can be rewritten as

$$q = \frac{n_4^{\text{eff}} |A_0|^2}{n_2 U_0^2}, \tag{2.103}$$

where $|A_0|$ and U_0 are real and represent the peak value of the field amplitude in real and normalized units respectively. Now the quintic nonlinear index n_4^{eff} is written in the convenient form

$$n_4^{\text{eff}} = -\frac{n_2}{A_{\text{sat}}^2}, \tag{2.104}$$

where A_{sat} is a real constant saturation amplitude. With this definition, the total refractive index seen by a scalar field is

$$n = n_0 + n_2 \left[1 - \frac{|A|^2}{A_{\text{sat}}^2} \right] |A|^2, \quad (2.105)$$

such that, for a peak field amplitude $|A_0| = \sqrt{\sigma} A_{\text{sat}}$, the induced nonlinear index will be reduced from the purely Kerr value of n_2 to $[1 - \sigma]n_2$, where $\sigma = |A_0|^2 / A_{\text{sat}}^2$ is the saturation parameter. The definition 2.104 is used so that the form of equation 2.105 is consistent with the first two terms of the Taylor expansion of the two-level system saturation [214].

Including only the cubic and quintic contributions as shown in equation 2.105 cannot result in true saturation. When $n_2 > 0$ and $n_4^{\text{eff}} < 0$ as in the situation here, the induced nonlinearity reaches a peak at $|A|^2 = A_{\text{sat}}^2/2$, then decreases and eventually becomes negative when $|A|^2 > A_{\text{sat}}^2$. This behavior is an indication that higher-order terms in the material polarization expansion may need to be included, but one experimental measurement of the induced nonlinear index in PTS at $\lambda_f = 1.064 \mu\text{m}$ shows that this roll-over into a negative contribution does indeed occur [147]. Nevertheless, when $\sigma < 0.5$, equation 2.105 mimics true saturation behavior.

Combining equations 2.103 and 2.104 and substituting into the eigenvalue equation 2.102 and noting that $U_0^2 / U_{\text{sat}}^2 = |A_0|^2 / A_{\text{sat}}^2$, results in the following ordinary differential eigenvalue equation

$$\frac{d^2 U}{d\rho^2} + \frac{1}{\rho} \frac{dU}{d\rho} + 2 \left[1 - \frac{U^2}{U_{\text{sat}}^2} \right] U^3 - \beta U = 0. \quad (2.106)$$

Figure 2.17 illustrates a few fundamental eigenmode solutions of equation 2.106 calculated as a function of peak normalized amplitude U_0 using the normalized saturation amplitude $U_{\text{sat}} = \sqrt{16.0}$. The heavy solid line ($U_0 = 1$) corresponds to the

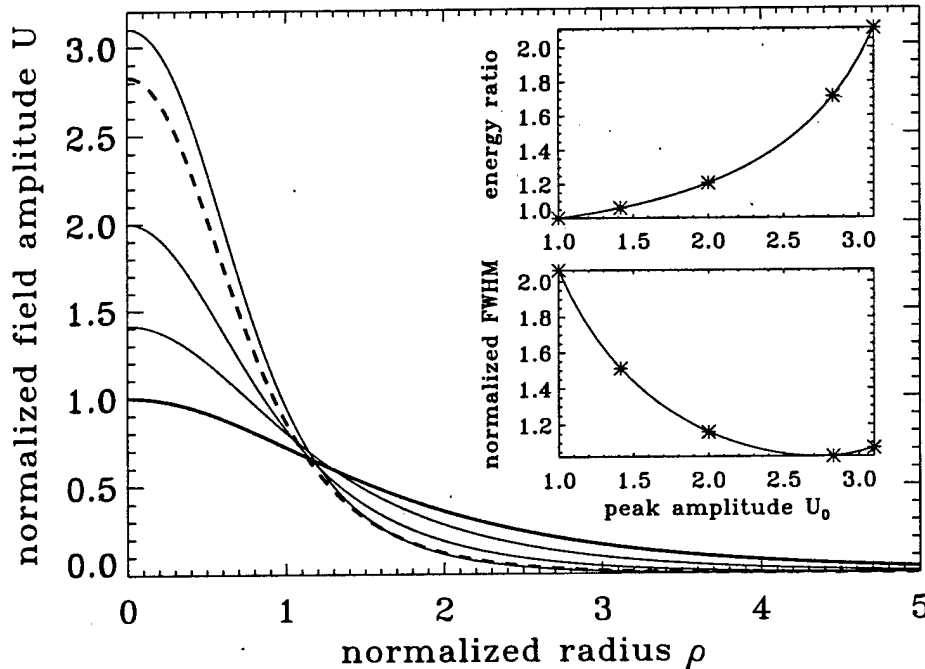


Figure 2.17: Fundamental eigenmodes of equation 2.106 plotted as a function of peak normalized amplitude. Here, the normalized saturation amplitude $U_{\text{sat}} = 4$, and $\sigma = 0.0625, 0.125, 0.25, 0.5$, and 0.6 for $U_0 = 1.0, 1.41, 2.0, 2.83$, and 3.1 , respectively. The inset plots show energy ratio (top) and normalized FWHM (bottom) as a function of U_0 , with asterisks denoting the positions of the eigenmodes in the main figure.

saturation parameter $\sigma = 0.0625$, while the heavy dashed line ($U_0 = 2.83$) corresponds to $\sigma = 0.5$. The top inset shows the energy ratio as a function of peak amplitude when the $\sigma = 0.0625$ eigenmode is used as the signal spatio-temporal wave. When using $\sigma = 0.5$ for the pump, the energy ratio is 1.712. The energy ratio increases with increasing σ , but when $\sigma > 0.5$, the

induced index as given by equation 2.105 begins to decrease from the maximum when $\sigma = 0.5$. The lower inset shows the variation in intensity full-width at half-maximum (FWHM). The FWHM initially decreases with σ , reaches a minimum near $\sigma = 0.5$, then increases as $\sigma \rightarrow 1$.

Figure 2.18 plots the $\sigma = 0.5$ fundamental eigenmode (dashed curve) of equation 2.106 along with $\sigma = 0$ eigenmodes (no saturation, solid curves) of peak normalized amplitudes of $U_0 = 2.83$ and $U_0 = 1.94$. The non-saturated $U_0 = 1.94$ eigenmode

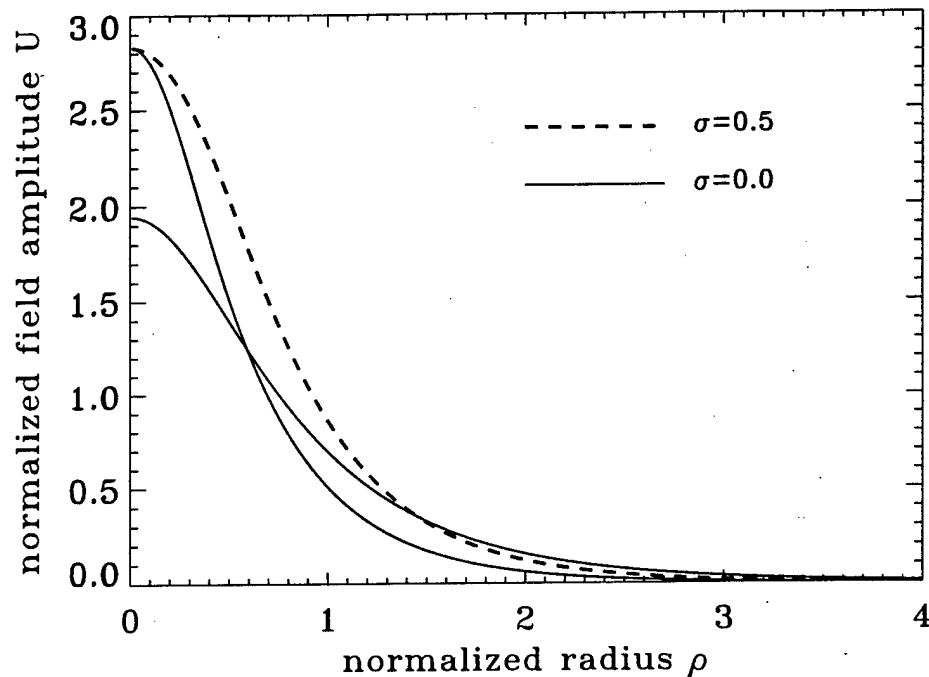


Figure 2.18: Numerically computed eigenmodes of equation 2.106 for the saturation parameter $\sigma = 0.5$ and normalized peak amplitude $U_0 = 2.83$ (dashed curves) and for the saturation parameter $\sigma = 0$ with $U_0 = 2.83$ and $U_0 = 1.94$ (solid curves). The effect of saturation is to broaden the width of the eigenmode near the peak where the saturation has its greatest effect. The space/time-bandwidth products are 0.312 for the saturated eigenmode and 0.242 for the non-saturated eigenmodes.

is a scaled version of the $U_0 = 2.83$ non-saturated eigenmode with the same normalized half-width at half-maximum (HWHM) of 0.5135 as the saturated $\sigma = 0.5$ eigenmode, and will be used in Chapter 6 to compare the effects of the higher-order temporal terms on the pump with and without saturation in section 6.1. Note that for the same HWHM, the saturated eigenmode will have a larger value of U_0 than the non-saturated one. The normalized FWHM of the intensity profiles of these eigenmodes is 1.027, while the FWHM of the $\sigma = 0.0625$ signal eigenmode is 2.063.

Stability of the cubic-quintic eigenmodes can be examined by the evaluation of the stability parameter [209]

$$S = \frac{\partial P}{\partial \beta}, \quad (2.107)$$

where P is the integration over the eigenmode profile and represents power in the 2-D spatial case and energy in the 3-D spatio-temporal case. Stability is ensured when $S > 0$, which is true for (1+1)-D NLS propagation, otherwise, propagation is unstable. For (2+1)-D propagation, $S = 0$, and for (3+1)-D, $S < 0$ [2]. The first condition is the result of the fact that the power of the 2-D eigenmode is constant with respect to width and eigenvalue [169], such that an increase in width, which reduces diffraction/dispersion, is exactly compensated by nonlinearity, and a decrease in width, which enhances nonlinearity, is exactly compensated by diffraction/dispersion. The second condition results from the fact that the energy of the 3-D eigenmode decreases with decreasing width and increasing eigenvalue [2], which is opposite to the 1-D case, such that an increase in width results in diffraction/dispersion overtaking nonlinearity, and a decrease in width results in nonlinearity overtaking diffraction/dispersion.

Figure 2.19 plots the eigenmode power (top) and stability parameter (bottom) versus eigenvalue for the (2+1)-D cubic-quintic NLS equation 2.106. It is shown that, like the 1-D case, the power increases with increasing eigenvalue (decreasing

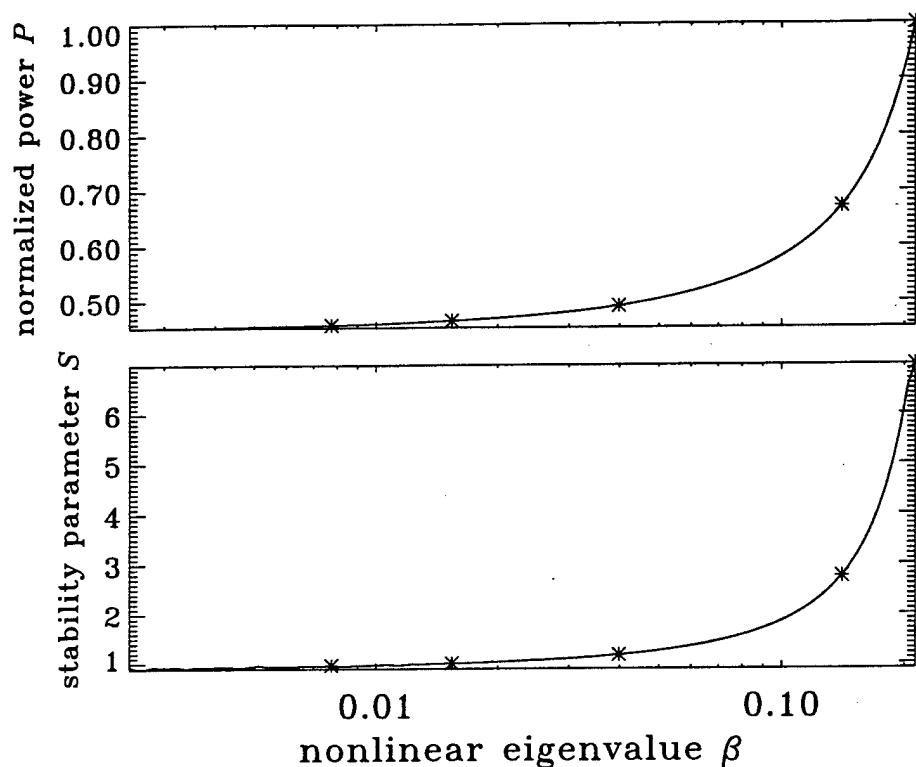


Figure 2.19: Plot of power and stability versus nonlinear eigenvalue, indicating that quintic saturation results in stable (2+1)-D eigenmodes. The asterisks denote the positions of the fundamental eigenmodes of Figure 2.17.

width), indicating stability, and S is always positive. The figure also indicates that eigenmodes with greater β (i.e. greater saturation) are more stable than eigenmodes with small β . The stability parameter asymptotically approaches zero with decreasing β which is expected since $\sigma \rightarrow 0$ as well.

Stable propagation of the $\sigma = 0.5$ cubic-quintic eigenmode is verified by numerical simulation of equation 2.98, as shown in Figure 2.20. The width parameter is chosen as $w_0 = 39.6 \mu\text{m}$, which results in intensity spatial FWHM of $40.7 \mu\text{m}$, and temporal FWHM of 16.5 fs. The propagation distance is $15 Z_0$, over which the spatio-temporal solitary wave is unchanged. The same stabilized solitary-wave behavior occurs in the fully 3-D case as well [2]. The confocal distance $Z_0 = 0.559 k_0 w_0^2 = 5.14 \text{ mm}$, and is calculated numerically. For the simulation, the parameters of fused silica (see Appendix D) are used at $\lambda_f = 1.55 \mu\text{m}$. The first section of Chapter 6 studies the effects of the higher-order terms of equation 2.99 on stability of this $\sigma = 0.5$ cubic-quintic pump solitary wave eigenmode, and shows that stabilization does occur, albeit with some spatio-temporal broadening due to Raman scattering.

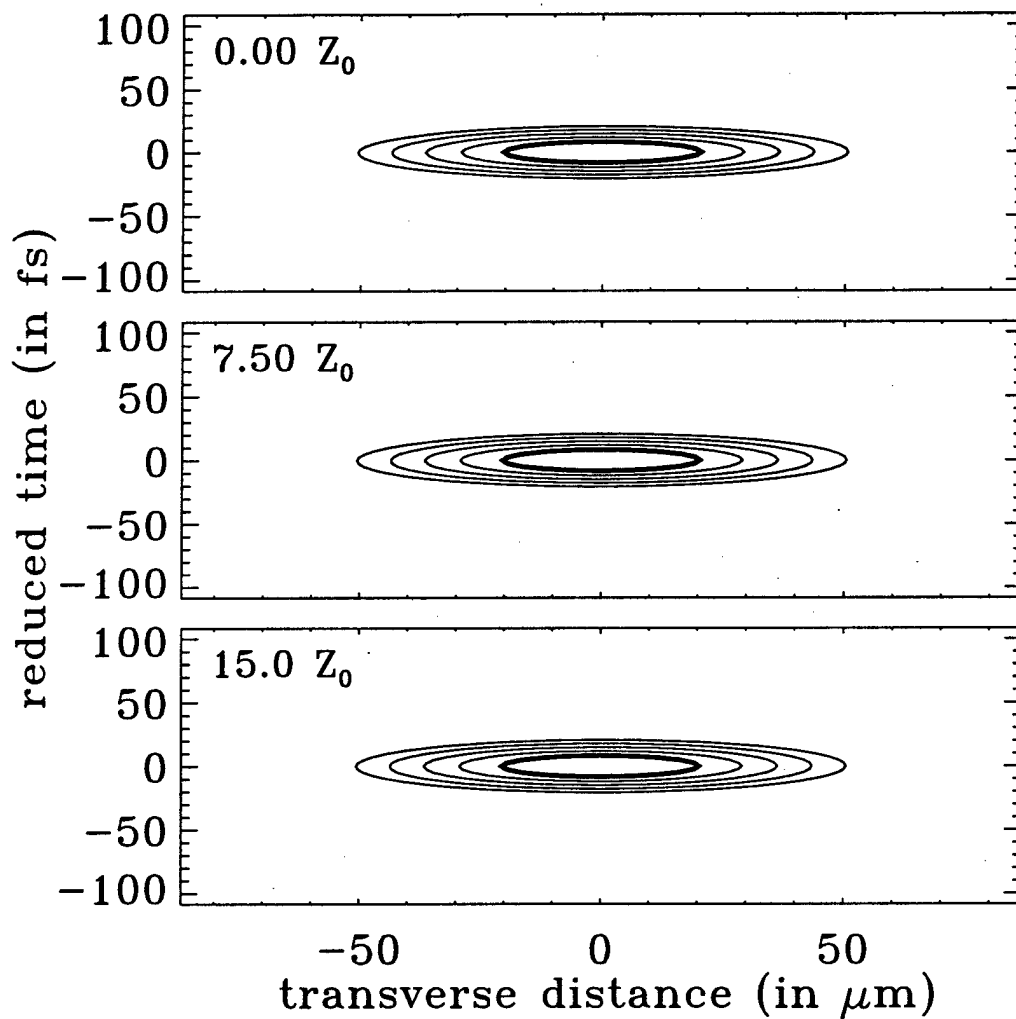


Figure 2.20: Stable propagation of the $\sigma = 0.5$ eigenmode of the cubic-quintic NLS equation without the effects of the higher-order terms. The spatio-temporal wave of transverse FWHM $40.7 \mu\text{m}$ and temporal FWHM 16.5 fs remains unchanged after propagating 15 confocal distances. The contours are at -3 dB intervals relative to the peak intensity in each frame.

Chapter 5

1-D Spatial Soliton Logic Gates

Interactions among optical solitons offer the promise of ultrafast all-optical switching and logic [2, 62, 72]. Many interactions have been studied theoretically and experimentally in recent years using both spatial and temporal solitons, but most of these geometries are only useful in single-stage optical switching. The focus of this thesis is on the implementation of cascadable soliton-based logic gates, because a good logic device can also be used as a switching device (a complete logic family can implement arbitrary switching functionality), but not vice versa since a switching device does not possess the requisite properties of logic level restoration and fanout with large noise-margins.

The types of interactions of interest here are those in which the presence of one or more solitons, called signals, affect the propagation of another, potentially more energetic, soliton called the pump, which is always present, and is analogous to the power supply in electronic gates. The pump soliton propagates by itself the length of the gate and passes through to the output to provide the high state of the device. Only this undeviated pump is used to switch subsequent stages, providing true three-terminal input-output isolation. When present, a signal soliton interacts with the pump, altering the propagation of the pump by inducing a spatial (or temporal) shift, angle (frequency) change, polarization rotation, etc., such that discrimination at the output is possible, thus providing the low output state of the device. In the case of one signal soliton, this class of interaction produces an inversion operation which can be cascaded to implement logically-complete, multi-input NOR gates, as originally suggested for temporal soliton interactions [66]. This thesis focuses on analogous spatial interactions [2, 58] in order to achieve high-density, low-latency switching in enhanced nonlinear media with simple output state discrimination.

Spatial and temporal interactions are analogous in the sense that spatial displacement is equivalent to temporal displacement, and angle (spatial frequency) is equivalent to temporal frequency. For a logic gate based upon a spatial interaction, it is much simpler to detect a displacement than an angle change, since the displacement can be discriminated using a spatial aperture in order to implement amplitude keyed logic. Note though, that an angle change can be mapped into a displacement using a Fourier lens, but this requires linear resolvability for high contrast between the high and low output states, which may be difficult to obtain. Conversely, for a temporal interaction, a temporal displacement is more difficult to detect than a spectral shift. Three terminal operation and amplitude keyed logic (for electronic detection) can simultaneously be obtained in the first case by using an ultrafast optical time gate based on nonlinear correlation to pass only the undisplaced pump. If the logic is time-shift keyed however, then timing will discriminate between a high and low state, but the logic state is not directly accessible to electronics. A spectral shift can be differentiated by using a passive frequency-selective optical element such as a prism, grating or bandpass/reject filter, but high contrast operation also requires that the spectral shift be resolvable, which, like a resolvable angular shift, is difficult to obtain, and usually precludes the simultaneous achievement of gain.

There are two ways considered here in which the signal can alter the propagation of the pump. The first is that the interaction can result in a spatial (temporal) shift for each soliton with no change in angle (temporal frequency). This shift of the pump is then independent of the gate length as long as the length is greater than that required for the interaction to occur. This case is less interesting because if the shift is not spatially (temporally) resolvable, then high contrast cannot be obtained. The other possibility is that the interaction produces an angle (temporal frequency) change in which the subsequent spatial (temporal) shift depends on the gate length. Since the pump soliton continues to propagate without broadening after the interaction, even in the situation in which the angle (temporal frequency) change is not linearly resolvable, arbitrarily large spatial (temporal) shift can be obtained by simply choosing a sufficiently long gate length, within the constraints placed by absorption. The lever arm nature of the temporal soliton dragging and angular deflection (i.e. spatial dragging) gates was discussed in sections 1.5.1 and 1.5.2, respectively.

This chapter studies interactions between 1-D spatial solitons in detail, and serves as a starting point for the study of logic gates based on spatial interactions between 2-D spatio-temporal solitary waves. Characterization of the interactions between spatial solitons is a much less demanding analytical and computational task, but still leads to significant insight into the multi-

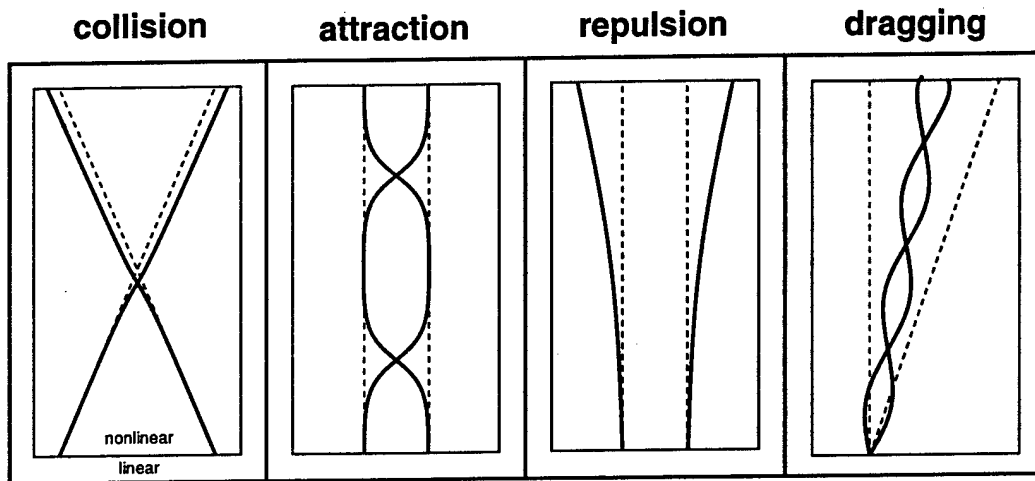


Figure 5.1: Illustration of the important spatial soliton interaction geometries. Each interaction is shown contained within a nonlinear medium, which itself is surrounded by a linear medium. The dashed lines indicate the soliton paths in the case of no interaction.

dimensional case. Spatial interactions are studied in this thesis because angular deflection leads to arbitrarily large spatial shift due to non-diffracting propagation and allows for the implementation of true three-terminal, amplitude keyed logic with the use of a simple spatial aperture designed to pass the undeflected pump, but block the pump after interaction with the signal. The analogous temporal interaction requires an ultrafast nonlinear time gating mechanism in order to provide these features.

Section 5.1 covers the basic soliton interaction geometries for both same-polarized (phase-sensitive) and orthogonally-polarized (phase-insensitive) inputs. Even though small-signal gain and fanout may be obtained for the phase-sensitive interactions, gate operation may be destroyed by small variation of the relative phase between the inputs and logic level restoration cannot be fully achieved because the phase is not restored at the output except under unrealistic circumstances. As a result, section 5.2 studies in detail the phase-insensitive collision and dragging interactions between orthogonally-polarized inputs. Based on the notion of a threshold contrast, the operation of these logic gates is studied under a wide range of parameters. The effects of material absorption on the operation of these gates is studied in section 5.3, and figures-of-merit are defined for use in evaluating low-loss nonlinear materials for soliton switching applications. Finally, section 5.4 considers logic gate properties in more detail, calculating the gate transfer function and numerically demonstrating cascable operation.

The main conclusions of this chapter are that, of the general class of three-terminal and restoring angular deflection logic gates, spatial soliton dragging logic gates have better performance characteristics than the spatial collision gates, and should provide cascable fanouts of 2 or greater with large noise margin. In addition, these studies identify material systems in which the gates and systems could be implemented and justify the pursuit of experimental investigation of these gates for ultrafast logic applications.

5.1 Soliton Interactions

Optical soliton interaction has been known since the first theoretical treatment of the (1+1)-D spatial optical nonlinear Schrödinger (NLS) equation via the inverse-scattering transform [73]. It wasn't until much later that these interaction forces were studied [74, 75]. The latter study [75] used the exact two-soliton solution [73] to determine that the interaction forces exponentially decrease with increasing soliton separation and sinusoidally depend on the relative phase. The motivation for the study was to quantify the temporal displacements (timing jitter) that might be suffered in a soliton communications system due to the various types of interactions that occur. These interactions, which are detrimental for communications [76-78], can be useful for switching, and are presented in the following sections.

The most common interactions that can produce this behavior are: collision, attraction/repulsion, and trapping/dragging, and are illustrated schematically in Figure 5.1. Typically, the collision interaction between two same-polarized solitons produces a spatial shift in the propagation of each soliton after collision. This interaction is a symmetric one in that the attractive forces between the solitons are balanced both before and after the point of complete overlap, and the solitons emerge unchanged (except for the spatial shift) due to the infinite number of conserved quantities of the NLS equation. Therefore, there can be no net change in propagation angle. A related collision occurs between solitons of orthogonal polarizations, and is described by

the coupled, or vector, NLS equations. In this case, the system of equations is not integrable (meaning that there are only a finite number of conserved quantities, the most fundamental being energy and momentum), and the collision can result in a permanent angle change, as desired. Soliton attraction results in a periodic exchange in the lateral position of two solitons, such that each soliton returns to its original spatial position after one period of oscillation. The period of oscillation depends on the size of the solitons and their initial separation, and the interaction results in neither a permanent spatial shift nor angle change. Repulsion occurs with the same initial geometry as attraction, but with solitons of the same polarization being π out of phase, rather than in-phase. This interaction does produce a permanent angle change, and the amount of change depends on the soliton widths and initial separation. Both attraction and repulsion interactions occur in the case of orthogonally polarized solitons, with the most useful interaction being a special case of collision at very small angles. Finally, spatial trapping/dragging is the only asymmetric interaction in that the point of soliton overlap is at the beginning of the nonlinear medium. This means that there is no interaction force before the point of overlap, and the force after the point of overlap is unbalanced, leading to a permanent angle change. These interactions will be discussed in more detail in this chapter, but only two - phase-insensitive collision and dragging between solitons of orthogonal polarizations - which produce angular deflection as discussed in section 1.5.2, are found suitable for application in logic. These two interactions are studied further in this chapter.

In order that the logic gate provide large signal gain (or fanout in cascaded systems), the pump soliton must carry more power through the output aperture than is present in some defined signal level at the input which produces the low output state. The gain should be at least unity to avoid the use of external amplification, but a minimum gain of ~ 4 (neglecting absorption) is desirable to compensate for unavoidable system losses such as Fresnel reflections and material absorption, and to provide a fanout of 2, which is the minimum required in order to implement general-purpose logic. The requirement of large-signal gain with high contrast is what limits the choice of available interactions.

The following sections discuss the basic parameters used to describe the initial conditions of an interaction, followed by a discussion of the interactions in the situations when the signal and pump polarizations are parallel (phase-sensitive) and orthogonal (phase-insensitive).

5.1.1 Definition of Parameters

Since spatial interactions are of ultimate interest to this thesis, the parameter κ is defined as the spatial frequency separation between two $\text{sech}()$ profiles divided by the sum of their spectral intensity half-widths at half-maxima (HWHM), and is referred to as the normalized interaction angle. An analogous parameter can also be defined based on the temporal frequency separations. Therefore, for mutually incoherent but otherwise identical beams, $\kappa = 1$ is the Sparrow resolvability condition [79] for Gaussian beams such that in the far field of linear propagation, the beam intensity profiles will overlap at the half-power points, which is approximately true for $\text{sech}()$ profile beams as well. Additionally, between same-polarized solitons, κ is related to the amount of phase-sensitivity of the interaction. When $\kappa \ll 1$, the interaction is highly phase-sensitive and when $\kappa \gg 1$, the interaction is phase-insensitive.

Assuming that one beam propagates down the optical axis (i.e. the z -axis), the following Fourier transform relation holds [80]

$$\text{sech}\left(\frac{x}{w_1}\right) \longleftrightarrow \frac{\pi w_1}{2} \text{sech}\left(\frac{\pi w_1 k_x}{2}\right). \quad (5.1)$$

The spectral intensity full-width at half-maximum (FWHM) is given by the condition

$$\frac{\pi w_1 \Delta k_x|_{w_1}}{2} = 1.7627 \Rightarrow \Delta k_x|_{w_1} = \frac{1.1222}{w_1}. \quad (5.2)$$

The other beam propagates at an angle θ to the optical axis, and by the Fourier shift theorem,

$$\text{sech}\left(\frac{x}{w_2}\right) e^{i\delta k_x x} \longleftrightarrow \frac{\pi w_2}{2} \text{sech}\left(\frac{\pi w_2}{2} [k_x - \delta k_x]\right), \quad (5.3)$$

where $\delta k_x = k_0 \sin \theta$. Now, κ is evaluated

$$\kappa \equiv \frac{2\delta k_x}{\Delta k_x|_{w_1} + \Delta k_x|_{w_2}} = \frac{k_0 w_1 w_2 \sin \theta}{0.5611[w_1 + w_2]} = \frac{11.20 w_1 w_2 \sin \theta}{\lambda[w_1 + w_2]}. \quad (5.4)$$

Another useful relation gives the transverse spatial frequency offset in terms of the normalized interaction angle κ

$$\delta k_x = \frac{0.5611[w_1 + w_2]\kappa}{w_1 w_2}. \quad (5.5)$$

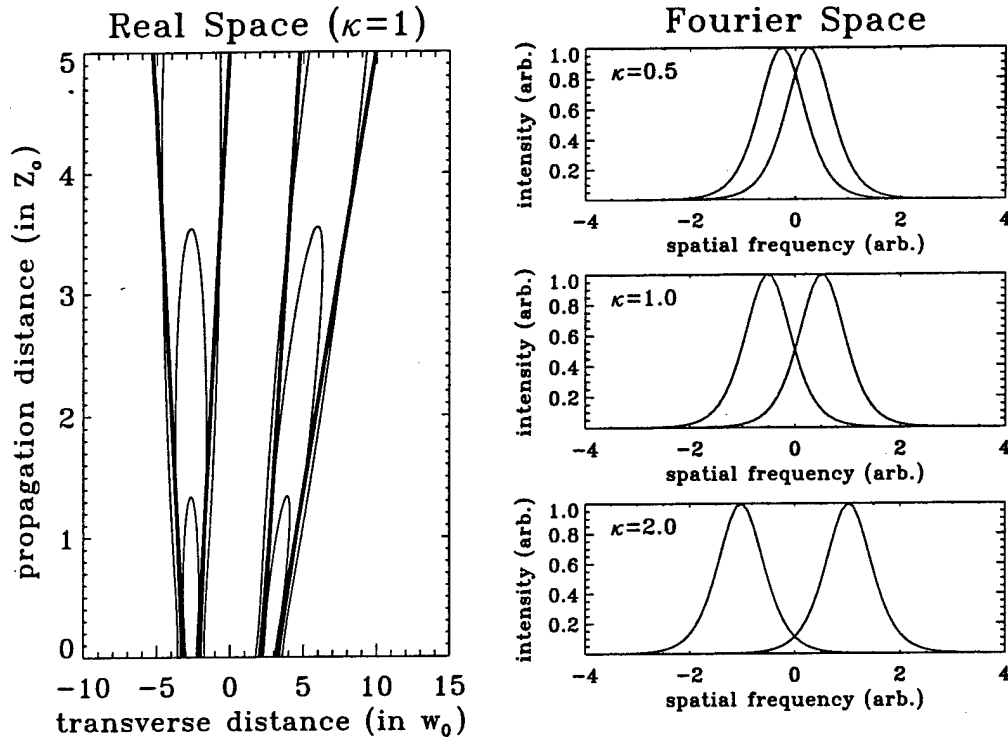


Figure 5.2: Illustration of the normalized interaction angle κ , as given by the spectral separation divided by the sum of the spectral HWHM. The left plot shows the $\kappa = 1$ condition in linear real space propagation where the beam on the left is propagating down the optical axis and the beam on the right is propagating with normalized angle $\kappa = 1$. The heavy lines indicate the spatial FWHM of the respective beams and the contours are spaced at 3 dB intervals relative to the initial peak intensity. The right-hand plots show the spectra of two beams for different values of κ . When $\kappa = 1$, the beams are at the resolvability condition (i.e. the spectra overlap at their half-power points).

For the 1-D spatial soliton interactions studied in this chapter, $w_1 = w_0$ is the signal soliton width and $w_2 = w_0/r$ is the pump soliton width where r is the initial power ratio.

Figure 5.2 illustrates the physical, linear, interpretations of the definition of κ in both real and Fourier space. In physical space, the $\kappa = 1$ condition indicates that the rate of separation between two beams is about the same as the rate of diffraction. Thus, the amount of overlap is constant with distance. In Fourier space, $\kappa < 1$ indicates large spectral overlap between two beams (i.e. not resolvable), and $\kappa > 1$ denotes small overlap (i.e. resolvable).

The confocal distance Z_0 was discussed in section 2.3.2, but is defined here as well for completeness. Here, for sech() shaped beams, $Z_0 = \pi^2 w_0^2 / \lambda = \pi k_0 w_0^2 / 2$ is twice the Rayleigh range, which is the distance over which the FWHM intensity pattern increases by a factor of $\sqrt{2}$. In the case of a Gaussian beam, the confocal distance is $k_0 w_0^2$. The confocal distance for a sech() profile was determined numerically¹ through linear beam propagation.

Finally, some metric must be defined in order to evaluate the potential for a soliton interaction to form the basis for a useful logic gate. A "threshold" contrast metric is presently defined as the power of the input (fundamental) signal divided by the power of the deviated pump that exits the aperture. This metric depends on the pump power, signal power, gate length, aperture size, and interaction angle, and is only used initially to provide a comparison among the different interactions and in later sections to locate the optimal interaction parameters, such as the choice of normalized interaction angle κ for a given gate length and aperture size. With this measure, threshold contrast greater than unity indicates that sufficient deviation has occurred such that the power exiting the aperture is less than the input signal power; contrast much greater than unity (i.e. 10 or greater) is an indication that a restoring logic gate can be implemented for the given interaction parameters, as discussed later. The choice of reference input signal level, which determines the unity threshold condition, is somewhat arbitrary. The two main options are the power of the signal soliton, or the fraction of that energy that is required to reach some acceptable low output level of the device. For simplicity, the fundamental signal soliton power will be used. The ramifications of this choice are discussed in section 5.2. A more precise definition of contrast will be employed later, based on the calculation of the gate

¹The actual numerically-obtained coefficient is not exactly π^2 . This coefficient is used instead for simplicity.

transfer function in section 5.4.

5.1.2 Phase-Sensitive Interactions

The first group of interactions is between solitons in the same polarization state and is described by the (1+1)-D scalar NLS equation in either space 2.23 or time 2.52, reproduced here in spatial form

$$2ik_0 \frac{\partial A}{\partial z} + \frac{\partial^2 A}{\partial x^2} + 2k_0^2 \frac{n_2}{n_0} |A|^2 A = 0. \quad (5.6)$$

Note that for the studies of this section, the fully paraxial NLS equation is sufficient. In order to describe the interaction between two solitons, the initial field amplitude distribution is taken to be of the form

$$A = \frac{1}{k_0 w_0} \sqrt{\frac{n_0}{n_2}} \operatorname{sech} \left(\frac{x + \zeta/2}{w_0} \right) e^{i\delta k_x x/2} e^{i\phi} + \frac{r}{k_0 w_0} \sqrt{\frac{n_0}{n_2}} \operatorname{sech} \left(\frac{x - \zeta/2}{w_0/r} \right) e^{-i\delta k_x x/2}, \quad (5.7)$$

where w_0 is the width parameter of the signal soliton, w_0/r is the width of the pump soliton, $\zeta = r_s w_0 [1 + 1/r]/2$ is the initial spatial separation defined so that the separation in terms of the sum of the spatial HWHM of the two beams is constant with r , δk_x is the initial relative transverse wavenumber as given by equation 5.5, and ϕ is the initial relative phase. For all simulations in this section, $r = 1$.

The phase-sensitivity of the interaction depends on the amount of angular or spectral overlap between the interacting solitons. When there is very little spectral overlap, such as the case for a large relative propagation angle or a large temporal frequency difference, even in regions of complete spatial or temporal overlap, the interaction will be phase-insensitive because the solitons pass through each other with large relative velocity. This also means that the interaction is very weak, though, and subsequently ineffective for a logic gate. Therefore, in order for the phase-sensitive interactions to perform high-contrast logic operations with gain, the solitons must have significant spectral overlap. The result is that the interaction necessarily depends on the relative phase between the two solitons, such that the contrast of the gate also depends on the relative phase, which may be difficult to control in an experimental situation. Because of the strong phase-dependence, this group of interactions is not useful for robust logic, but is presented mainly for historical and pedagogical reasons and for comparison with the phase-insensitive interactions of section 5.1.3.

Soliton Collision

Collision between optical NLS solitons is conservative in the sense that the solitons pass through each other without change except for a small spatial or temporal shift [75]. Theoretical studies of multi-soliton collisions have been performed [78,81] for the purposes of quantifying timing jitter in WDM soliton systems. Experimental demonstrations in optics have been carried-out for temporal soliton collision in fiber [82,83], and dark spatial soliton collision in self-defocusing liquid media [84,85]. The latter experiments prompted a study of using the dark soliton collisions for general-purpose logic [72]. The collision interaction between bright spatial solitons has also been suggested as a mechanism for a photonic switch [86], although no experimental realization of that device has been presented to date.

The collision interaction occurs when two spatial (temporal) solitons, propagating at different angles (with different frequencies leading to different group velocities), symmetrically collide in a nonlinear medium, producing a spatial (temporal) shift in the propagation of each soliton, with no permanent change in propagation angle (frequency). This interaction is illustrated in Figure 5.3 for spatial solitons. After the collision, when the solitons are well separated, the contrast is independent of gate length, because the initial propagation angles are restored and the spatial shift assumes a fixed value. Therefore, shorter gate lengths result from greater angles of collision. One problem with the collision interaction, though, is that the spatial shift decreases with increasing angle of collision, meaning that long gate lengths are required in order to obtain high contrast operation by using small collision angles.

This scaling can be quantified by considering the analytic expression for the spatial shift after collision [78]

$$\Delta x = \pm w_0 \ln \left(1 + \frac{4}{w_0^2 \delta k_x^2} \right), \quad (5.8)$$

where Δx is the displacement of a single soliton (the \pm indicates that the solitons shift in opposite directions) and δk_x is the total spatial frequency separation between the interacting solitons. Setting the gate length to dZ_0 and the initial spatial separation to

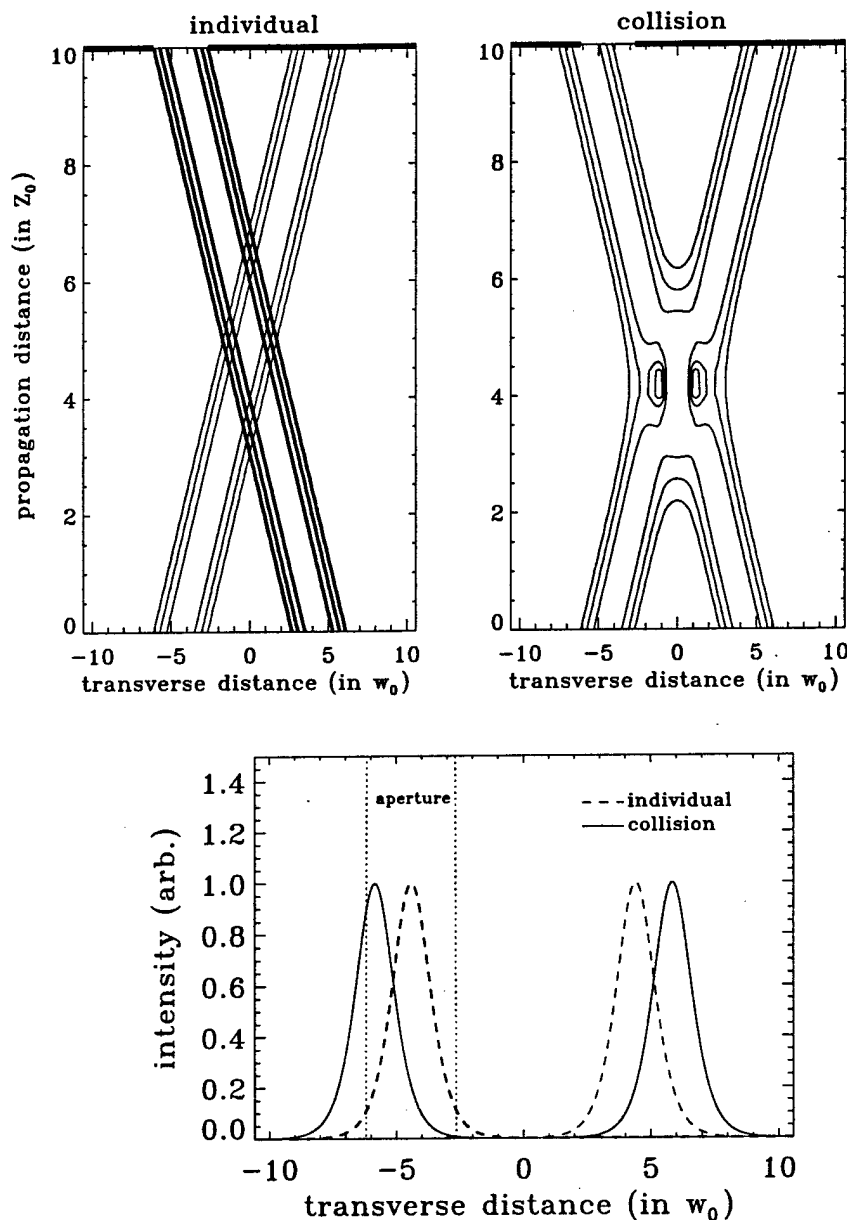


Figure 5.3: Collision interaction between spatial solitons. The initial separation is 5.0 FWHM ($= 8.81 w_0$) and the propagation distance is $10 Z_0$, where $Z_0 = \pi^2 w_0^2 / \lambda$. Each soliton propagates at an angle given by the tangent of the initial separation divided by the gate length, and the total normalized angle is $\kappa \approx 1$. The bottom plot indicates the initial positions of the solitons (dashed curves) and the final positions after collision (solid curves). If there were no interaction, the final positions would be the reverse of the initial positions. The threshold contrast of this gate is 1.5.

$r_s w_0$, such that $\delta k_x = 2k_0 \tan(r_s w_0 / d Z_0) \approx 2r_s k_0 w_0 / d Z_0$ (so that $\kappa \approx 1.135 r_s / d$), the fractional shift of each soliton is written

$$\frac{\Delta x}{w_0} = \pm \ln \left(1 + \frac{\pi^2 d^2}{4 r_s^2} \right). \quad (5.9)$$

If $r_a w_0$ equals the aperture width, then the desired minimum spatial shift Δx is also $r_a w_0$. Note that $r_s \geq r_a$. Finally, an expression for the minimum gate length in terms of the initial separation r_s and aperture width parameter r_a is obtained

$$d Z_0 \geq \frac{2 r_s}{\pi} \sqrt{e^{r_a} - 1} Z_0, \quad (5.10)$$

such that for an initial separation of 5.0 FWHM = 8.81 w_0 and aperture width 2.0 FWHM = 3.5 w_0 , the gate length must be greater than 32 Z_0 . This choice of aperture size allows 94.2% of the power of the undeviated pump to pass. Therefore, for a 2 FWHM spatial shift, only 2.9% of the pump power exits the aperture, and threshold contrast of 34.5 is obtained.

In addition to long gate length, another problem that results from small-angle collision is that, even though the spatial shift can be greater, the phase-sensitivity is stronger since the overlap of the angular spectra becomes significant. This is evident in the next section on attraction/repulsion, which are the limiting cases of small angle collision. As will be shown in section 5.1.3, collision between solitons of orthogonal polarizations gets around this problem of phase sensitivity and is one of the interactions studied later in this chapter.

Soliton Attraction/Repulsion

The first experimental study of soliton interactions in optical fiber demonstrated the attraction/repulsion interaction between temporal solitons of the same frequency [87]. Experimental demonstrations of spatial soliton attraction and repulsion have also been reported in liquid CS₂ cells [88] and glass slab waveguides [89]. An all-optical switch based on the spatial interaction has been proposed and analyzed [90], but not experimentally realized.

In this interaction geometry, two spatial (temporal) solitons launched along parallel paths (with the same frequency) separated by a few beam widths (pulse durations) attract, repel, or exhibit more complicated behavior, depending on the relative phase. This interaction is strongly phase-sensitive because the angular (frequency) spectra of the interacting solitons overlap completely. As mentioned previously, this is the limiting case of small-angle collision. Greater initial spatial (temporal) overlap between the two solitons leads to a stronger interaction, but coherent beam combining can result in random transmission when the sources are not phase-locked to within a fraction of a radian, as they must be in order that the nonlinear interaction is repeatable. The use of a beamsplitter to combine the two beams results in a 3 dB loss in the regions of non-overlap since the solitons initially propagate with the same angle (frequency) and are of the same polarization. This 3 dB loss is unacceptable for optical logic because it ultimately reduces the gate fanout by a factor of 2, with the loss taken not at the output of a gate, but at the inputs to subsequent gates.

When the solitons are in-phase, they constructively interfere and attract, as shown in Figure 5.4, due to the increase in nonlinear index in the regions of overlap. This attraction causes the solitons to coalesce, then pass through each other and exchange initial positions. The interaction repeats indefinitely, each time the solitons exchange positions after passing through each other. The output of the gate is set to the first of the infinite number of distances at which the solitons completely overlap, which is where the soliton positions are changed the most, resulting in the maximum contrast. Note that, like the case for soliton collision, the contrast is determined by the interaction, not the gate length (except that the gate length must be appropriately chosen to coincide to one of the distances of complete overlap). The plot on the bottom of Figure 5.4 shows the soliton profiles when there is no interaction (with either one of the solitons present) and at the first point of complete overlap when the interaction occurs (both of the solitons present). It is evident then that higher contrast can be obtained with larger initial separation, but since the interaction is much weaker in this case, the gate length will necessarily increase.

Adapting previous results [74] on phase-sensitive interactions into the notation of this thesis, the normalized separation between the solitons as a function of propagation distance for the attraction interaction is

$$r_s(z) - r_s(0) = 2 \ln \left| \cos \left(2e^{-r_s(0)/2} z / k_0 w_0^2 \right) \right|, \quad (5.11)$$

where $r_s(0) w_0$ is the initial separation in real units. The gate length $z = d Z_0$ is set at the first position of complete overlap. At this position $r_s(z) = 0$ and straightforward manipulations yield

$$d = \frac{\cos^{-1} \left(e^{-r_s(0)/2} \right)}{\pi} e^{r_s(0)/2}. \quad (5.12)$$

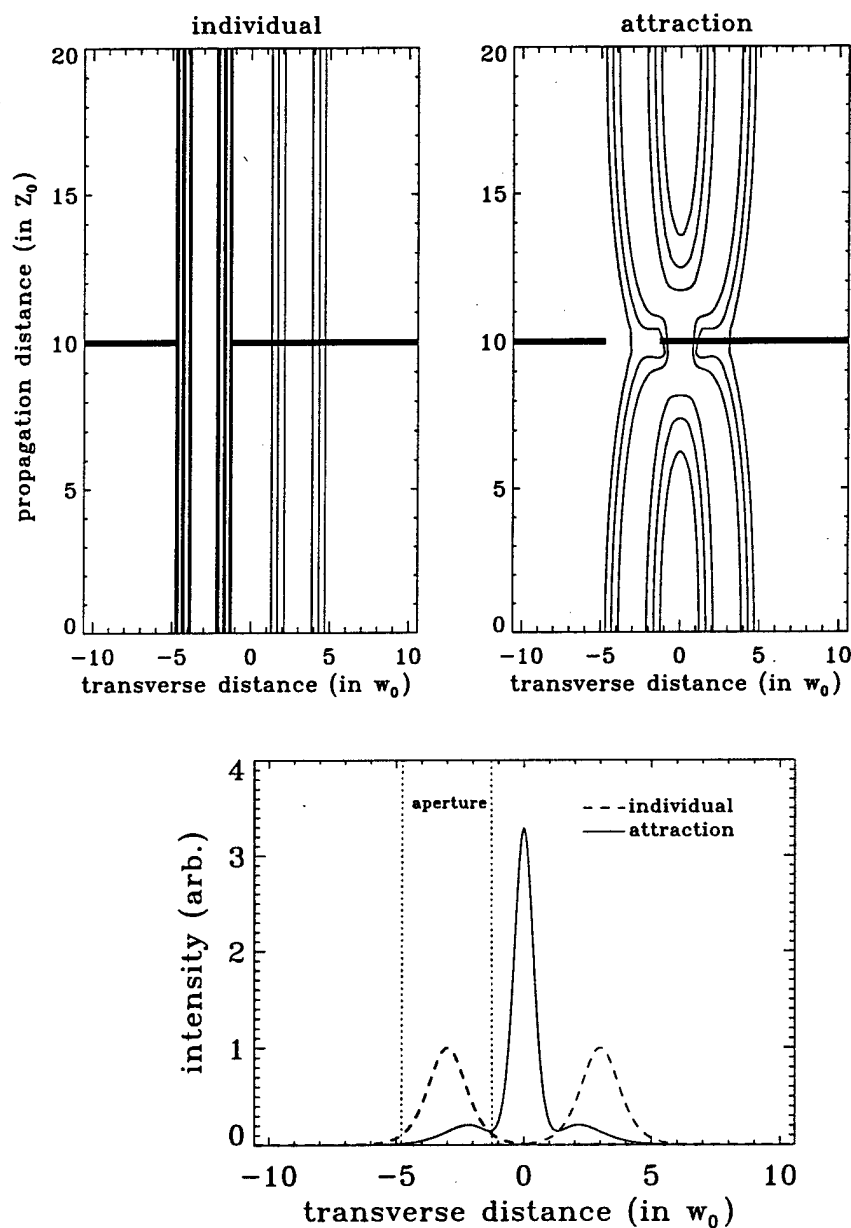


Figure 5.4: Attraction interaction between spatial solitons. The initial separation is 3.4 FWHM ($= 6.0 w_0$) and the propagation distance is $20 Z_0$, which is one-half of the interaction period. The solitons are initially in-phase and propagate along a parallel path, but constructively interfere and attract due to the increase in nonlinear index in the overlap regions. The plot on the bottom indicates the initial positions of the solitons (dashed curves) and the positions after $10 Z_0$ where the solitons overlap (solid curves). This interaction has a period of $40 Z_0$, and the threshold contrast of the gate is 4.8.

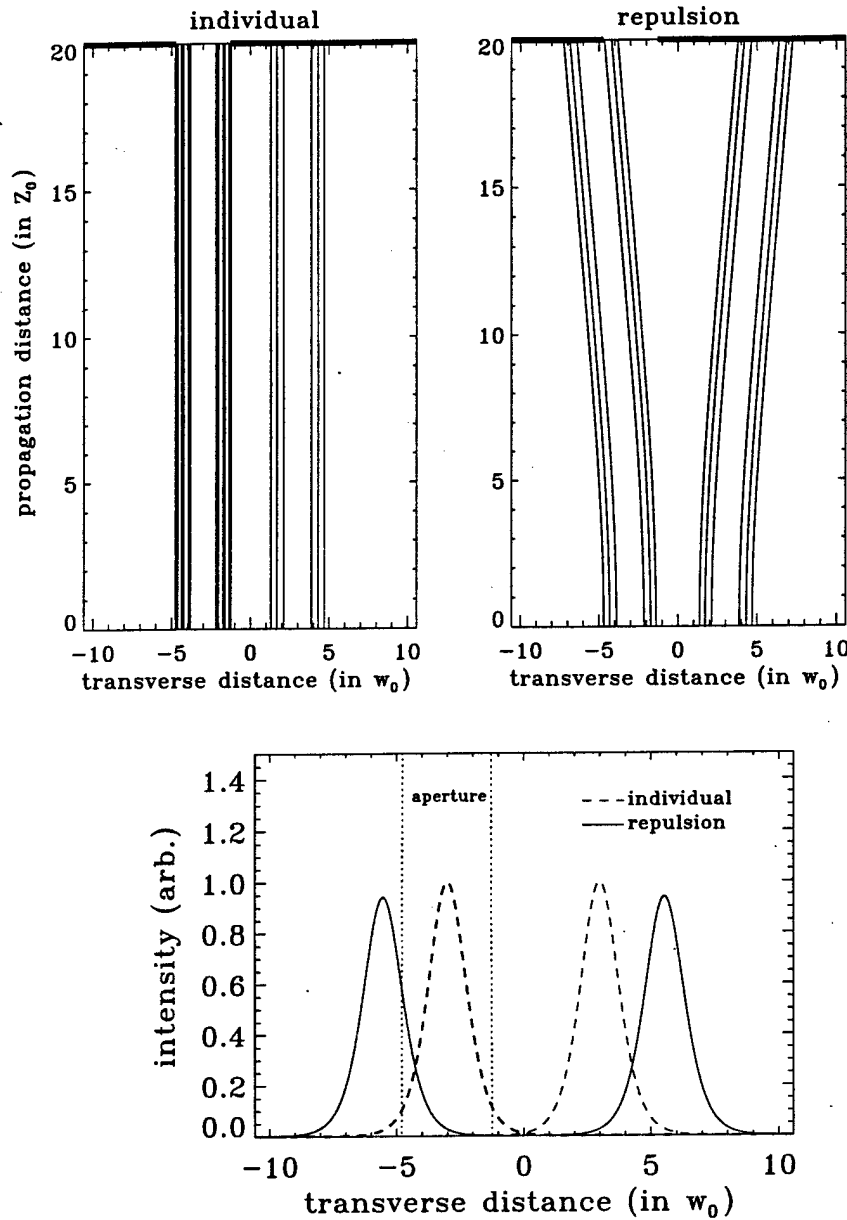


Figure 5.5: Repulsion interaction between spatial solitons. The initial separation is 3.4 FWHM ($= 6.0 w_0$) and the propagation distance is $20 Z_0$. The solitons are initially π out-of-phase and propagate along a parallel path, but destructively interfere and repel due to the decrease in nonlinear index in the overlap regions. The plot on the bottom indicates the initial positions of the solitons (dashed curves) and the positions after $20 Z_0$ where the solitons repelled (solid curves). The threshold contrast of this gate is 5.5.

For initial separations $r_s(0) \gg 1$, this distance can be approximated as

$$d \approx \frac{e^{r_s(0)/2}}{2}. \quad (5.13)$$

As an example, to obtain a position shift equal to the aperture width $r_a w_0$, the initial separation is $r_s(0) = 2r_a$. If $r_a = 3.5$, then the minimum gate length for the interaction is $d = 17$.

When the solitons have π relative phase, they destructively interfere and repel, as shown in Figure 5.5, due to the decrease in nonlinear index in the regions of overlap. Unlike the previous case, this interaction results in a persistent change in propagation angle of each soliton, such that the angle change is constant after the solitons are well separated. In this case, the threshold contrast increases with longer gate length. Alternatively, for the same gate length, the contrast increases with decreased initial

separation between the solitons, but again the use of a beamsplitter results in 3 dB coupling loss in the regions where the beams do not overlap. In the limit of complete overlap, the interaction degenerates to linear beamsplitter logic, in which the relative phase between the light entering the input ports determines the amount of light exiting each of the two output ports.

Similar to the attraction case, an analytical expression for the separation during the repulsion interaction is given by [74]

$$r_s(z) = r_s(0) + 2 \ln \left(\cosh \left(\pi d e^{-r_s(0)/2} \right) \right). \quad (5.14)$$

The gate length required to obtain a separation $r_s(d) - r_s(0)$ is given by

$$d = \frac{e^{r_s(0)/2}}{\pi} \cosh^{-1} \left(e^{[r_s(d) - r_s(0)]/2} \right), \quad (5.15)$$

which can be reduced to

$$d \approx \frac{e^{r_s(0)/2}}{2\pi} [r_s(d) - r_s(0)] \quad (5.16)$$

when $r_s(d) \gg r_s(0)$. Again, if the initial separation is $r_s(0) = 2r_a$, then $r_s(d) = 4r_a$ leads to a well resolved spatial shift of either soliton and $d = 44$ when $r_a = 3.5$.

When the relative phase is a value other than 0 or π , other behaviors are possible [91]. For $\pi/2$ relative phase, power/energy is exchanged such that the soliton with the phase delay grows while the phase advanced soliton decays. The solitons also weakly repel each other. Once again, the solitons need large overlap in order for the interaction to produce a high-contrast switching operation, implying a 3 dB input coupling penalty as well.

By now it should be clear that this class of interaction is highly phase-dependent, requiring precise control and stability of the optical system for consistent operation. It has also been known in temporal soliton literature that the interaction (collision, attraction and repulsion) between temporal solitons of the same color can be significantly reduced by using solitons of unequal amplitudes [75, 92, 93], but threshold contrasts (and gate length scalings) similar to those discussed here for $r = 1$ can be obtained in the spatial analog situations by scaling the output aperture width with the pump width so that large-signal gain can be obtained. However, the phase-sensitivity presents a major practical difficulty.

Soliton Trapping

In spatial (temporal) collision, there is no net angle (frequency or group velocity) change, but a permanent spatial (temporal) shift. During the collision, though, there is a temporary change in angle (frequency), such that the change is largest at the point when the solitons completely overlap and returns to zero when the solitons are well separated. The integral of this angle (velocity) change gives the spatial (temporal) shift. When the symmetry of this interaction is broken by absorption/amplification [78, 94] or initial overlap [64, 95], some angle (velocity) change remains after the solitons separate. This effect is detrimental to WDM communications [96], but useful for optical switching [63, 68] because the permanent angle (velocity) change can be integrated over long distances to provide large spatial (temporal) shift, as discussed in section 1.5.2.

Trapping of initially overlapping spatial solitons propagating in different directions (the spatial analog of temporal trapping using two different color solitons in a fiber [63]) has been demonstrated experimentally [68, 97]. This interaction is very similar to the dragging geometry discussed in the next section, but is highly phase dependent. Trapping between initially-overlapping in-phase spatial solitons is shown in Figure 5.6 where the solitons propagate away from each other at angles given by $\kappa = \pm 0.5$, for a total angular separation of $\kappa = 1$. The solitons mutually attract, and, when the propagation angles are not too great, as in Figure 5.6, form a bound state. The net result is a change in propagation angle such that the contrast depends on the gate length (the effective trapping distance is essentially zero since the solitons initially overlap).

When the overlapping components of the two spatial frequency spectra are not in-phase, more complicated behavior results. As the relative phase is changed, interference fringes move through the overlapping intensity region such that, for $\kappa = 1$, a π relative phase shift places a single dark fringe at the center of the overlap region. Non-trapping behavior is shown in Figure 5.7, where the soliton initially propagating to the left is delayed in phase by 0.85π relative to the soliton initially propagating to the right. This value of phase was chosen for this particular gate length to illustrate an approximately worst-case effect. In this case, power is transferred to the left-propagating soliton as the solitons repel. The final positions at the output of the gate are nearly the same as would occur without the interaction. By changing the phase, the final position at the output can be moved around. This effect has been called soliton steering [63], and is detrimental for optical switching when phase cannot be precisely controlled.

For complete trapping between equal amplitude solitons, it is straightforward to calculate the gate length required for high-contrast switching. Assuming a total angle of propagation θ between the solitons, the required gate length for a shift of $r_a w_0$ is

$$d = \frac{r_a w_0}{\tan(\theta/2) Z_0}. \quad (5.17)$$

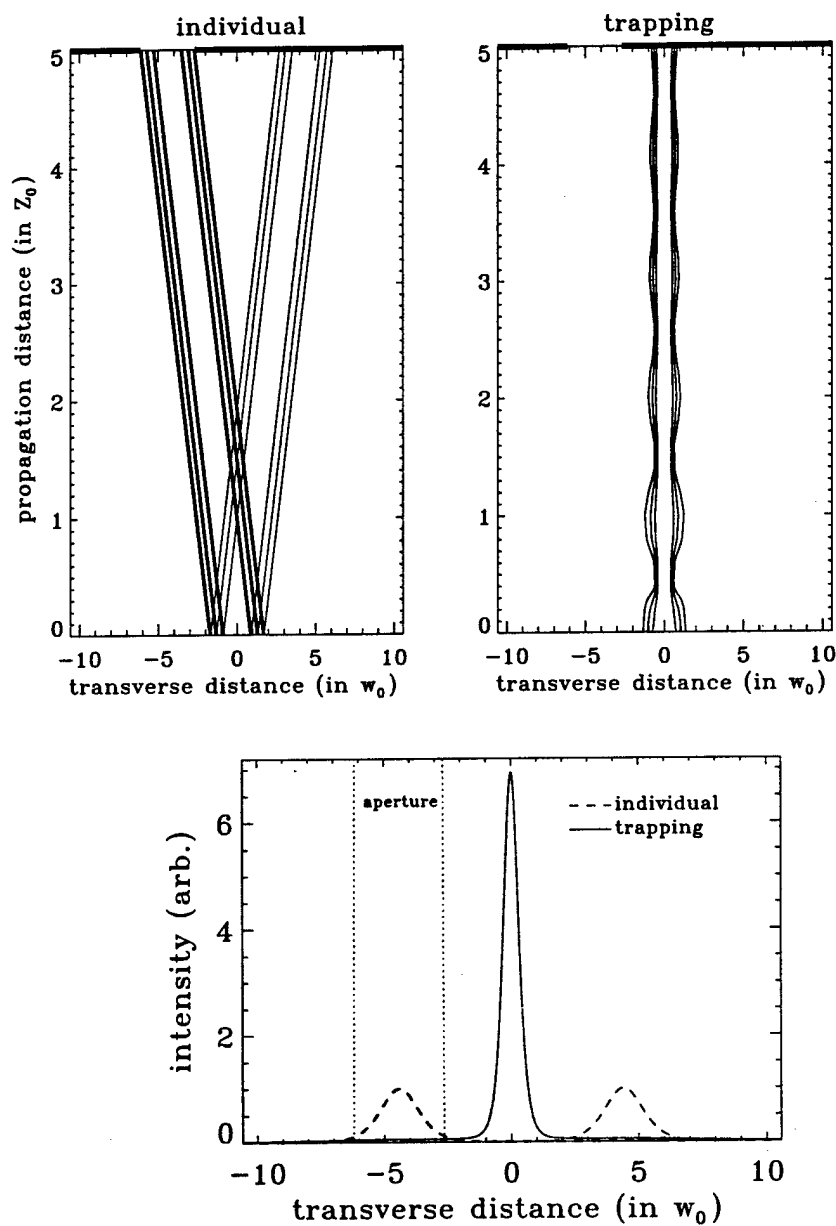


Figure 5.6: Spatial trapping interaction between initially overlapping in-phase spatial solitons. Each soliton propagates at an angle given by $\kappa = \pm 0.5$. The plot on the right indicates the final positions of the solitons both without interaction (dashed curves) and after trapping (solid curves). The threshold contrast of this gate of length $5 Z_0$ is 11.

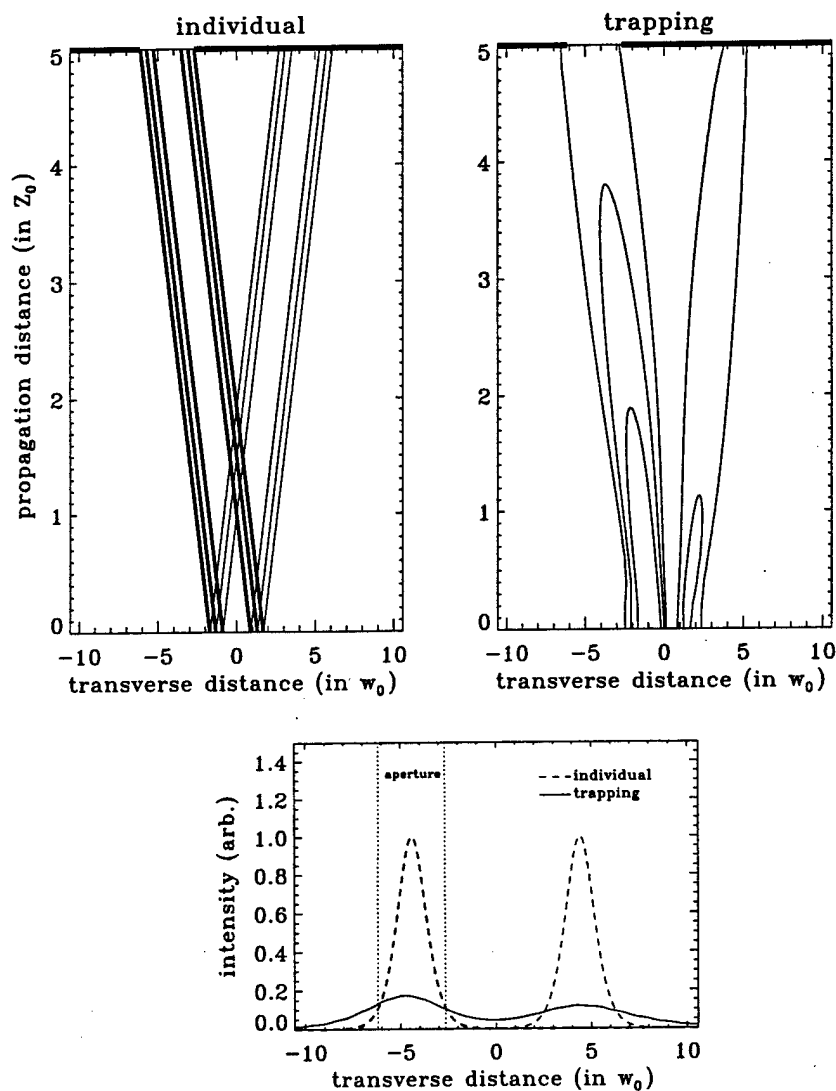


Figure 5.7: Spatial trapping interaction between initially overlapping out-of-phase spatial solitons (the left soliton lags in phase by 0.85π). Each soliton propagates at an angle given by $\kappa = \pm 0.5$. The plot on the right indicates the final positions of the solitons both without interaction (dashed curves) and after trapping (solid curves). The threshold contrast of this gate of length $5 Z_0$ is 3.8

For small θ , $\tan(\theta/2) \approx \sin(\theta/2) \approx \sin(\theta)/2$, and from section 5.1.3, $\sin(\theta)/2 = 0.1786\kappa\lambda/2w_0$. The gate length can now be written more simply as

$$d = \frac{1.135r_a}{\kappa}. \quad (5.18)$$

In the example of Figure 5.6, the minimum gate length required for a shift of $3.5 w_0$ is $4.0 Z_0$, when $\kappa = 1$ represents total normalized angle. The minimum gate length can be determined by knowing the maximum value of κ such that the solitons remain bound. Theoretical [64] and numerical/experimental [63] studies place this value at $\kappa \leq 2.7$ for equal amplitude solitons. This sets a lower bound on the gate length at around $1.5 Z_0$, considerably shorter than in the cases of collision and attraction/repulsion.

Numerical studies show that significant gain can be achieved with high contrast with the trapping interaction between in-phase solitons by scaling the aperture size with pump width, but also, with the worst-case phase difference, the contrast drops proportionally with the pump to signal power ratio, or approximately $1/r \ll 1$, rendering the gate unusable. Nevertheless, as will be verified later in the thesis for soliton dragging, it is expected that this interaction will be very useful in the phase-insensitive case of orthogonal polarizations.

5.1.3 Phase-Insensitive Interactions

In a computing or switching system consisting of a large number of gates with arbitrary interconnections and feedback paths, a fixed relative phase may be difficult if not impossible to maintain, therefore it is essential to reduce or completely eliminate phase-dependence. This can be accomplished by using orthogonally polarized solitons, such that, in a nonlinear Kerr material, the soliton interaction will be based only on phase-insensitive nonlinear cross-focusing (typically referred to as cross-phase modulation for temporal solitons) between their intensity profiles. Phase-dependence is completely eliminated in the case of pseudo-1D soliton interactions in bulk [98] or perfect fiber through the use of orthogonal circular polarizations. In the case of orthogonal linear polarizations, an additional phase-dependent nonlinear term (the vector four-wave mixing term) enters the coupled NLS equation. In this situation, phase dependence can be reduced by use of material or waveguide birefringence to mismatch the phase-dependent terms in the induced nonlinear material polarization. The latter is the method use in this thesis since the 1-D spatial and 2-D spatio-temporal solitons are confined in one dimension by a slab waveguide which admits orthogonal linearly-polarized TE and TM solutions, each propagating with a different phase velocity. This difference is due to the form birefringence of the waveguide, and can be tailored by variation of the waveguide parameters, as shown in Appendix D. Note that in the case of standard optical fiber, which has circular symmetry, the birefringence is very weak ($\Delta n \approx 10^{-6} - 10^{-5}$) and is due to micro bends and slight imperfections and residual stress due to manufacturing, but the very long interaction lengths still allow for the phase-dependent terms to be neglected in most situations.

Neglecting the phase-dependent nonlinear (i.e. the vector four-wave mixing) terms then, the interaction between orthogonally-polarized solitons is described by the coupled NLS equations, written here in spatial form,

$$\begin{aligned} 2ik_0 \frac{\partial A_x}{\partial z} + \frac{\partial^2 A_x}{\partial x^2} + 2k_0^2 \frac{n_2}{n_0} \left[|A_x|^2 + 2\Delta |A_y|^2 \right] A_x &= 0 \\ 2ik_0 \frac{\partial A_y}{\partial z} + \frac{\partial^2 A_y}{\partial x^2} + 2k_0^2 \frac{n_2}{n_0} \left[|A_y|^2 + 2\Delta |A_x|^2 \right] A_y &= 0, \end{aligned} \quad (5.19a)$$

where 2Δ represents the strength of the nonlinear coupling between the polarizations. This equation is the spatial analog of that used for temporal solitons in fiber [67, 99]. In isotropic Kerr media, $\Delta = 1/3$, which is the value used throughout this thesis. For this choice of Δ , the coupled equations are not integrable (in fact, they are only integrable when $2\Delta = 1$ [100]) and approximate analytical or numerical techniques must be used. The result is that some interactions are inelastic [101], meaning that the solitons do not necessarily retain their initial character after collision. As will be shown in the following sections, this behavior can be advantageous for optical switching. A consequence of the inelastic nature is that the solitons can be pulled apart by the interaction, resulting in "shadow" beams (pulses [71]) co-propagating with the stronger soliton involved. Note that in the case $2\Delta = 1$, the coupled equations are known as the Manakov equations which are integrable [100], resulting in interactions with similar behavior to those between same-polarized, in-phase solitons, but without the phase-dependence. The case $2\Delta = 1$ does not exactly correspond to any material symmetry class, but to a coincidental combination of the third-order nonlinear susceptibility components of cubic AlGaAs at sub-half bandgap [102].

In section 5.1.2, the point was argued that the strength of an interaction depends on the amount of spectral overlap between the solitons and resulted in phase-dependence due to scalar linear interference. In the case of interaction between solitons of orthogonal polarizations, though, there is no linear interference and the interaction depends only on the nonlinear coupling between the intensity profiles. It is expected that these interactions will also be stronger for $\kappa < 1$ because this implies that the transverse velocities are small, such that each soliton remains bound by the "potential well" [103] of the other soliton. This

potential well is created by cross-phase modulation (cross-focusing), and gives rise to an escape velocity [64, 104] which is the minimum (transverse and/or longitudinal) relative velocity necessary for one soliton to escape the potential well of the other.

The interactions studied in the previous section are now examined for the phase-insensitive case of orthogonally-polarized solitons. Here, the initial field amplitude distributions are taken to be of the form

$$A_x = \frac{r}{k_0 w_0} \sqrt{\frac{n_0}{n_2}} \operatorname{sech} \left(\frac{x - \zeta/2}{w_0/r} \right) e^{-i\delta k_x x/2} \quad (5.20a)$$

$$A_y = \frac{1}{k_0 w_0} \sqrt{\frac{n_0}{n_2}} \operatorname{sech} \left(\frac{x + \zeta/2}{w_0} \right) e^{i\delta k_x x/2}, \quad (5.20b)$$

where w_0 is the width parameter of the signal soliton, w_0/r is the width of the pump soliton, $\zeta = r_s w_0 [1 + 1/r]/2$ is the initial spatial separation, and δk_x is the initial relative transverse wavenumber as given by equation 5.5. Note that, even if the solitons are otherwise identical (i.e. $r = 1$), they are now distinguishable by polarization state in the absence of the vectorial four-wave mixing term, and any signal leakage through the aperture can be blocked by a linear polarization analyzer. The signal and pump can also be brought into coincidence through the use of a polarizing beamsplitter, without 3 dB input coupling loss.

Soliton Collision

For soliton collision in the integrable case of equation 5.19a when $2\Delta = 1$, the solitons pass through each other [100] simply producing a small spatial shift as in the same-polarized interaction, with no change in angle (temporal frequency). But when $2\Delta \neq 1$, more interesting behavior occurs, because the angles (frequencies) can change after collision [105]. A theoretical study of the collision between two orthogonally-polarized temporal solitons with $2\Delta = 2/3$ showed that reflection, annihilation, and fusion could result after the collision [101]. This study was concerned with the deleterious effects that the inelastic collision (resulting in annihilation and fusion) had on optical communications and switching. With the type of three-terminal logic gate studied here though, it does not matter what happens to the pump soliton as long as it doesn't pass through the aperture in the presence of the signal. Therefore, these inelastic interactions are potentially useful for three-terminal optical switching.

One result of this work was the derivation of a resonance condition which determined when the interaction is highly inelastic such that the solitons form a bound pair after collision [101]. This resonance condition depends on the group-velocity difference (angle for the present purposes) and relative amplitude of the two solitons. A later theoretical study [106] determined the permanent frequency shifts versus fiber birefringence after inelastic collision between equal-amplitude temporal solitons and noted that these frequency (angle) shifts become the temporal (spatial) shifts due to nonzero group-delay dispersion (paraxial diffraction). The same group numerically investigated the use of the inelastic collision between orthogonally-polarized spatial solitons as logic gates [107].

Figures 5.8 and 5.9 show some examples of these collision interactions using spatial solitons. In each case, the initial separation is $r_s = 5.29$. Collision between equal amplitude $r = 1$ solitons is shown in Figure 5.8. The top contour plots show (elastic) collision at large angles which behaves similarly to the (phase-sensitive) scalar collision. In the case of unequal amplitudes $r = 3$ as shown at the top of Figure 5.9, part of the weaker signal soliton breaks away after collision with the pump, resulting in an inelastic interaction. In either case though, gates based upon large angle collision still have lower contrast.

Small angle collision, on the other hand, can provide for high contrast logic gates, as shown at the bottom of the two figures. Small angle collision is termed slow collision in the temporal soliton literature [101]. In this case, nearly equal-amplitude solitons (Figure 5.8) will fuse into a bound orbiting pair while unequal amplitude solitons will mutually reflect (Figure 5.9). This reflection interaction is very useful for optical switching because it occurs between solitons of unequal amplitude (thus allowing gain) and allows for twice the transfer of momentum between the interacting solitons. As shown in Figure 5.9, these small angle collisions can provide the basis for very high contrast gates with gain, and will be studied later in this chapter for this purpose.

Note that, if desired, the initial separation can be much smaller for collision between orthogonally-polarized solitons versus phase-sensitive collision because the solitons are always resolvable at the output through the use of an analyzer.

Soliton Attraction/Repulsion

Parallel propagating, orthogonally-polarized solitons are the limiting case of small-angle collision; however, it is not clear that the phase-insensitive attraction/repulsion interaction has been studied in the literature as the basis for a logic gate. Like orthogonally-polarized collision, solitons with nearly the same amplitudes will attract, and with sufficiently different amplitudes they will reflect or repel. In addition, the same type of inelastic behavior occurs in this case such that the solitons do not necessarily exit the interaction intact.

Examples of attraction and repulsion are shown in Figure 5.10. Attraction between equal amplitude solitons is shown at the

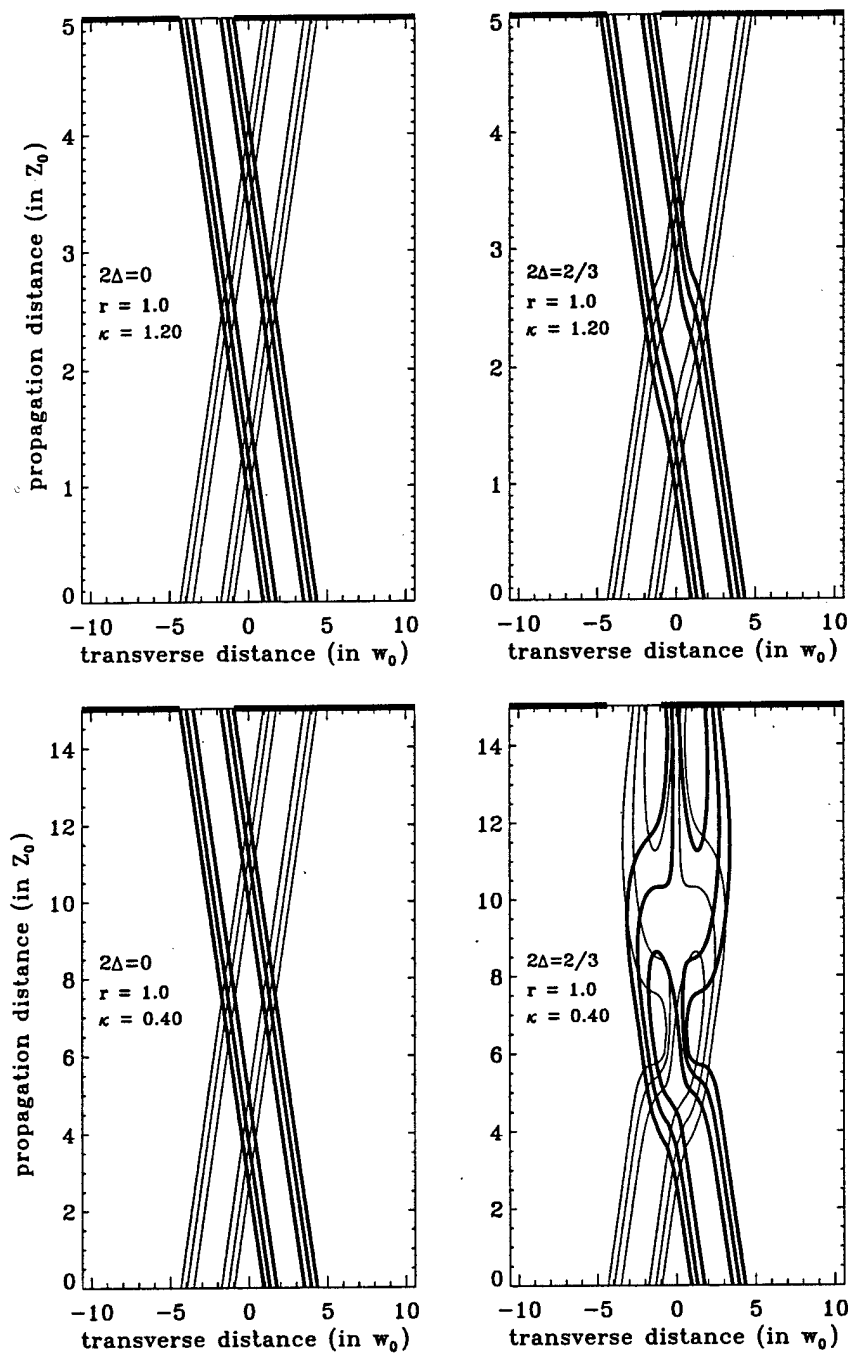


Figure 5.8: Collision interaction between orthogonally-polarized spatial solitons with equal amplitudes $r = 1$ and initial separation 3.0 FWHM of the signal ($= 5.29 w_0$). The top interaction has a gate length $5.0 Z_0$ and achieves a threshold contrast of 1.1 and the bottom fusion interaction has a gate length of $15 Z_0$ and achieves a threshold contrast of 23. Each soliton propagates at an angle given by the tangent of the initial separation divided by the gate length, and the value of $\kappa = 5.599 w_0 \sin \theta / \lambda$ listed in the figures corresponds to the total angle of propagation θ between the solitons.

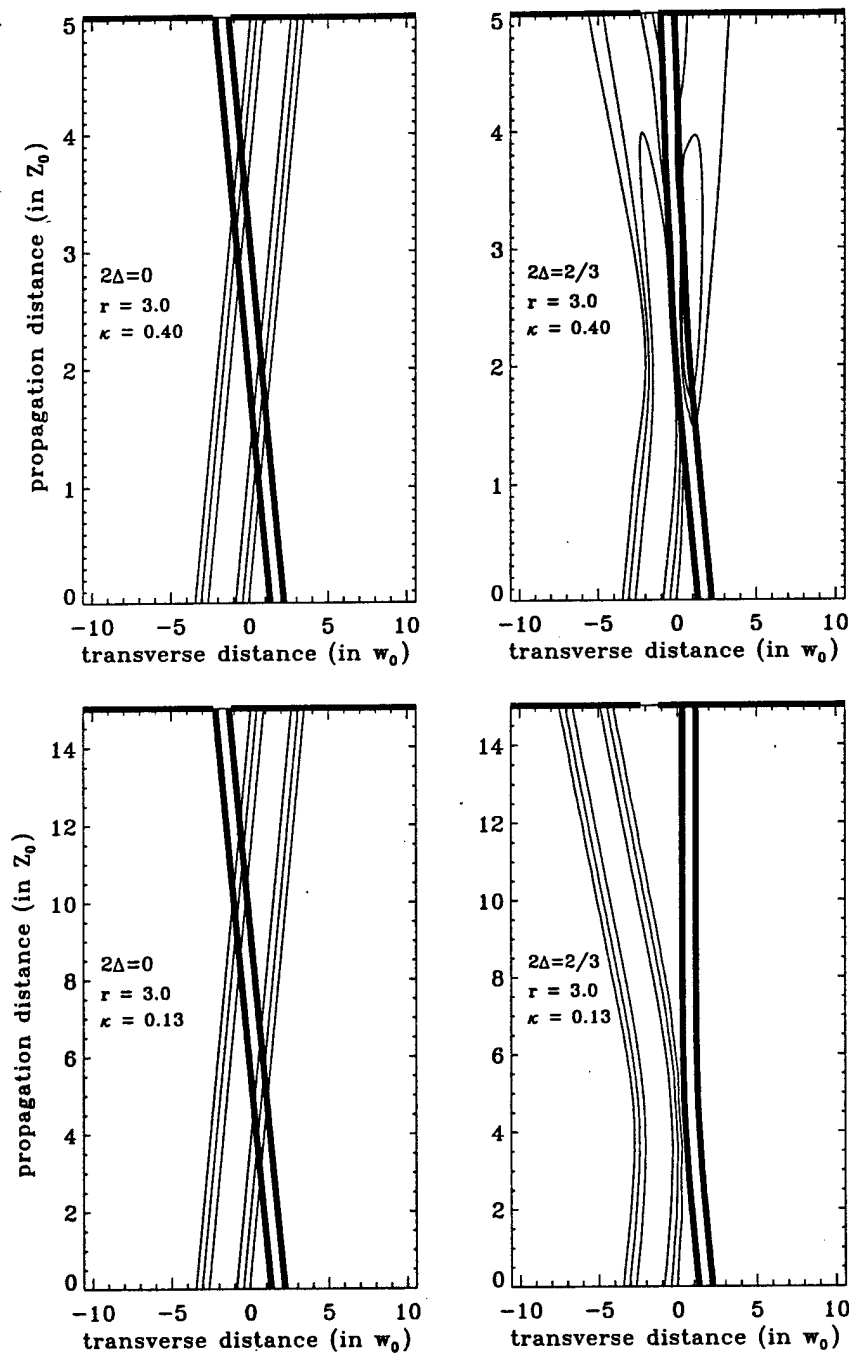


Figure 5.9: Collision interaction between orthogonally-polarized spatial solitons with unequal amplitudes $r = 3$ and initial separation $1.5 \text{ FWHM}[1 + 1/r]$, where $\text{FWHM} = 1.7627 w_0$. The top fusion interaction has a gate length $5.0 Z_0$ and achieves a threshold contrast of 9.6 and the bottom deflection interaction has a gate length of $15 Z_0$ and achieves a threshold contrast > 1000 . Each soliton propagates at an angle given by the tangent of the initial separation divided by the gate length, and the value of $\kappa = 11.20 w_0 \sin \theta / \lambda [r + 1]$ listed in the figures corresponds to the total angle of propagation θ between the solitons. The aperture width is given by $3.5 w_0 / r$.

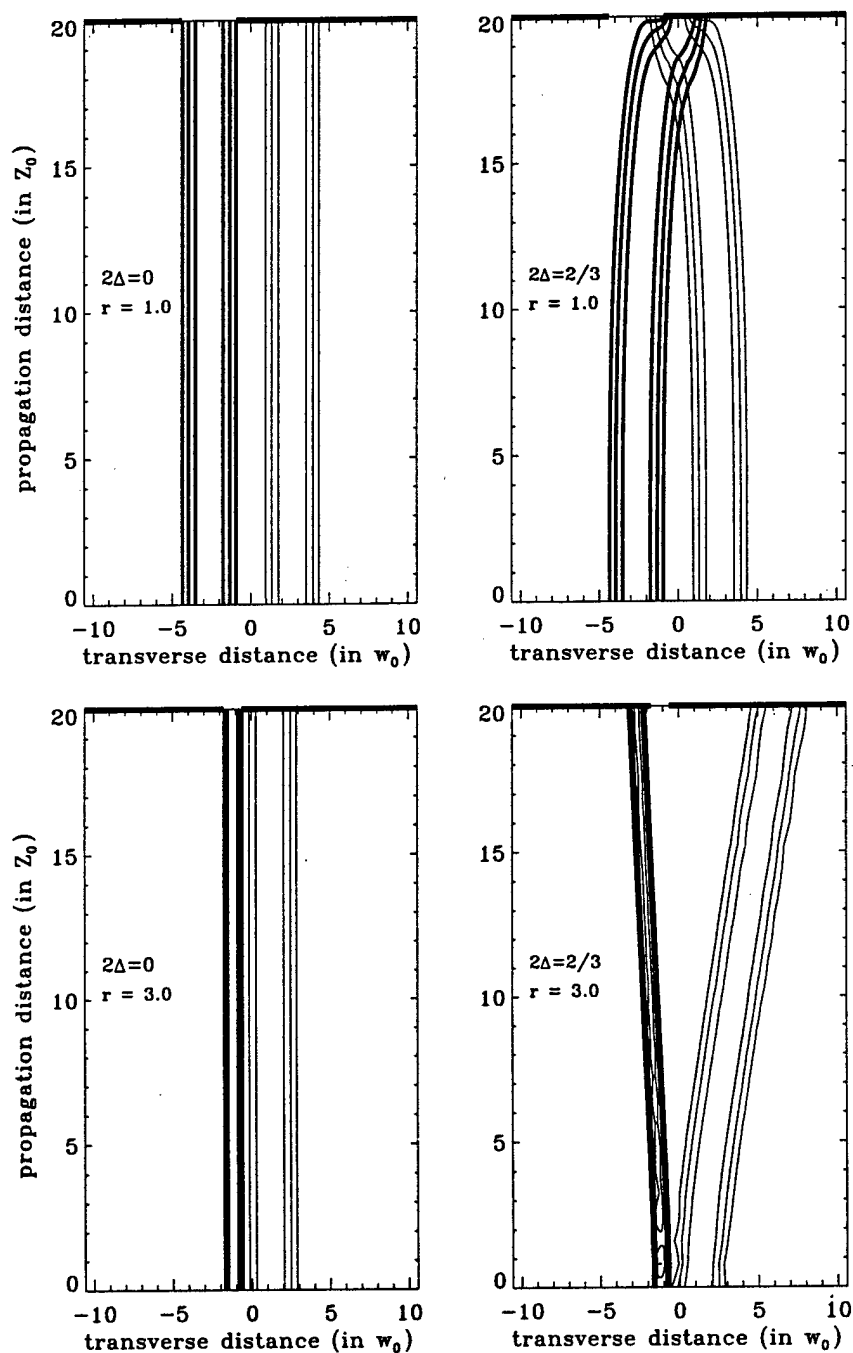


Figure 5.10: Attraction interaction between orthogonally-polarized spatial solitons with a gate length of $20 Z_0$. The top shows attraction between equal amplitude solitons with initial separation of 3.0 FWHM ($= 5.29 w_0$), achieving a threshold contrast of 9.8. The bottom shows repulsion between solitons where the pump power is $r = 3$ times the signal and the initial separation is $1.0 \text{ FWHM}[1 + 1/r]$ ($= 2.35 w_0$), achieving a threshold contrast of 69.

top, while repulsion between unequal amplitude solitons is shown at the bottom. The repulsion example shown is an inelastic interaction in that part of the signal breaks off and becomes bound to the pump. Repulsion becomes stronger as the initial separation is decreased from the well-separated condition, but then a point is reached where the interaction weakens. This can be understood intuitively by noting that when the solitons completely overlap at the input, symmetry dictates that no angular or spatial deviation of the center of mass is possible and the threshold contrast drops to $1/r$. Therefore, there must be an optimal separation that lies between the well-separated and overlapping conditions. The symmetry at complete overlap can be broken by allowing one or both of the solitons to initially propagate at an angle, which is the collision interaction with complete initial overlap. This interaction is known as spatial soliton dragging and is examined in the following section.

Soliton Dragging

Temporal soliton dragging is based upon the mechanism of trapping between orthogonally-polarized temporal solitons in weakly birefringent fiber [56, 57]. The original theoretical studies were concerned with using nonlinear attraction (cross-phase modulation) to overcome soliton pulse splitting due to group-velocity walkoff between the polarization modes, such that the partial soliton in each mode shifts its temporal frequency via cross-phase modulation induced chirp to compensate for the group-velocity birefringence. The induced chirp (leading to average spectral shift) causes the soliton along the slow axis to speed up (spectral up-shift in the AGDD regime) and the soliton along the fast axis to slow down (spectral down-shift). Above a critical energy, these partial solitons form an orbiting bound state. Experimental studies [65] verified this trapping mechanism by launching identical and temporally overlapping pulses (which individually had less energy than required to form a temporal soliton) polarized along the principle axes of highly birefringent fiber, which can be accomplished by rotating the polarization of the launched soliton with respect to the fiber birefringence axes.

This experiment led the way to using the trapping mechanism for optical switching [62], in which a polarization insensitive spectral bandpass filter was used to pass one pulse in either pure linear polarization state and block the double-lobed spectra of mutually-trapped pulses to implement the binary exclusive OR (XOR) operation. Since the output of the XOR gate can be in either of two linear polarization states, cascable operation is not possible without incurring a 3 dB loss to a 45 degree oriented polarizer. Orienting this output polarizer to the pump polarization, though, converts this gate into an inverter with a single linearly polarized output pulse representing the high state when the signal is absent, such that the pump passes through both the bandpass filter and polarizer. Gain is difficult to achieve with these gates because the spectral shift of the pump decreases (due to incomplete trapping) with increasing gain, leading to low contrast. The spectral shifts are a constant of the interaction and can be increased by using fiber with larger birefringence, but the maximum gain that can be obtained and still allow for complete trapping (i.e. maximum spectral shift), decreases [57].

Over long propagation distances, a slight temporal frequency change manifests itself as a large change in arrival time due to the fiber dispersive delay line, analogous to the operation of a lever arm. This observation led to the use of the timing shifts of the trapping interaction as the logical output state [59]. An experimental realization of a two-input NOR gate [59] launched a (normalized amplitude) $N = 1.5$, 132 pJ pump soliton along the fiber fast axis and two well-separated $N = 0.7$, 30 pJ signal pulses along the slow axis. If the pulses are separated too far apart, then sufficient overlap with the pump does not exist to induce dragging, and phase-sensitive operation results if the signal pulses are too close together and overlap. The presence of either or both signal pulses dragged the 300 fs pump by 2-5 ps in time over a distance of 75 m (25 soliton periods), realizing a gain of 4.5. This interaction became known as temporal dragging because the slow-axis pulse literally drags the fast-axis pulse in time, and has the capability of gain with high contrast and cascable operation [59] since only the undragged pump has the correct timing to serve as the signal input to the next gate and the signal pulses are blocked by the output analyzer. As mentioned previously, this is time-shift keyed logic and requires large temporal guard bands to achieve practical three-terminal operation.

The strong potential of the temporal dragging gate as a three-terminal logic device with gain and high contrast combined with the disadvantages of long gate lengths (many 10's of meters) and lack of spatial parallelism motivated the study of the spatial soliton dragging interaction [69]. The direct analogy consists of launching at normal incidence two initially overlapping spatial solitons into an anisotropic crystal cut at an angle with respect to a symmetry plane. In the case of an off-axis uniaxial crystal, the soliton in the ordinary polarization will propagate straight and the soliton in the extraordinary polarization will experience power flow walkoff away from the optic axis. These solitons will mutually attract and form a bound pair under the appropriate conditions of crystal cut (which determines walkoff angle), and soliton power ratio. Since the nonlinear cross phase modulation is between the intensity profiles, nearly the same behavior results by launching two initially overlapping solitons with tilted angles in isotropic media [58]. For an appropriate gate length, a large pump is dragged well over a beam diameter by a small signal and no longer passes through a spatial aperture. This revised geometry is more analogous to two-color soliton dragging in the temporal case. A three-terminal logic gate using spatial dragging should be easier to implement than temporal dragging because a simple spatial aperture will pass an undragged pump and block a dragged pump, instead of using ultrafast time gating. This spatial interaction has been experimentally demonstrated [108] with a gate length approximately

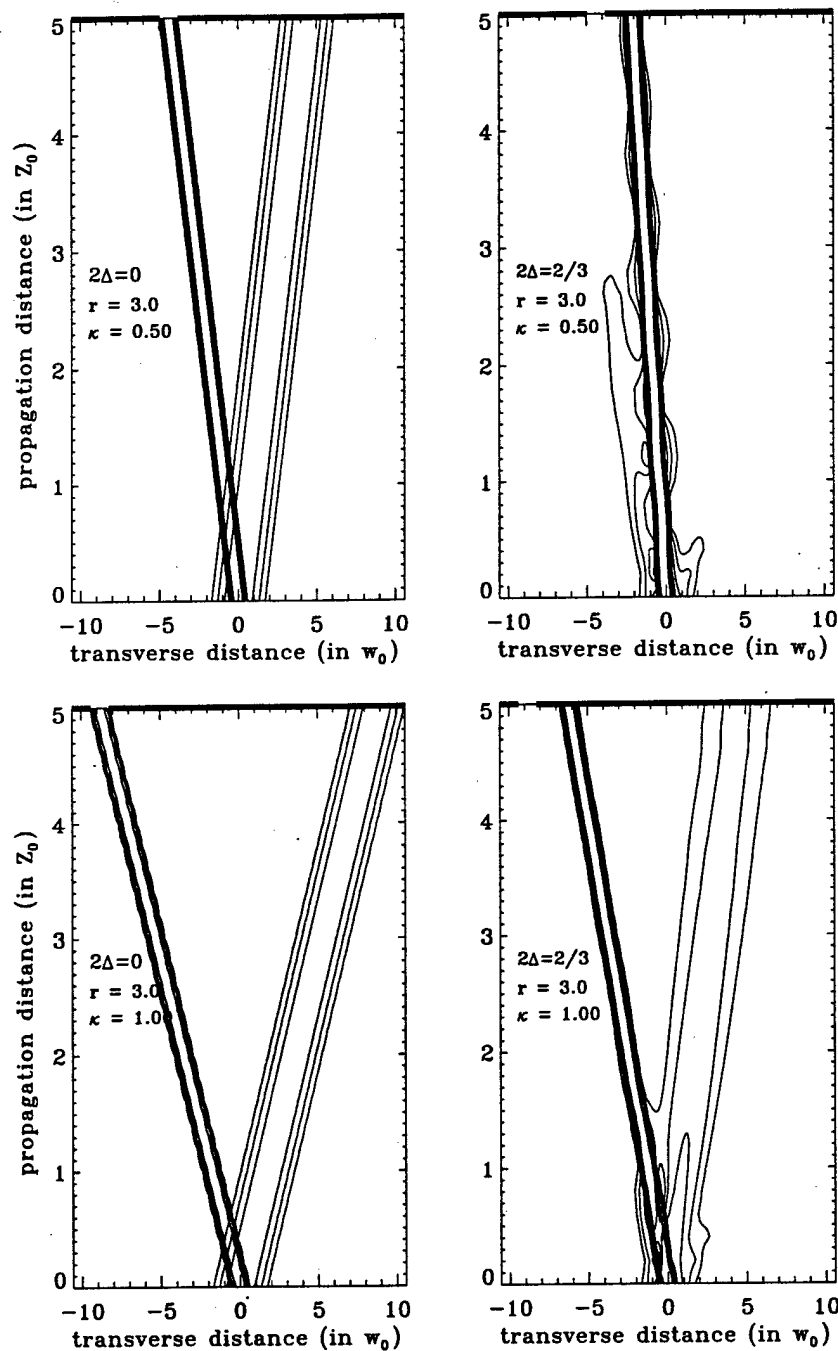


Figure 5.11: Dragging interaction between orthogonally-polarized spatial solitons of unequal amplitudes ($r = 3$). The initial total propagation angles between the solitons are given by $\kappa = 0.5$ (top) and $\kappa = 1.0$ (bottom) and the gate length is $5 Z_0$ in each case. Each soliton acquires an angle shift due to mutual trapping so that the final displaced positions depend on gate length. The threshold contrasts of these gates are > 1000 .

4 Z_0 , although the angular deflection achieved was very small due to experimental difficulties.

Figure 5.11 shows small (top) and large (bottom) angle spatial dragging. Note that now the solitons can be overlapped without loss at the input at small angles (less than resolvable) without violation of constant radiance. Small angle dragging $\kappa < 1$ (for nearly equal-amplitude solitons) typically results in complete trapping such that the interacting solitons form a bound orbiting pair (as shown) and propagate at the weighted mean angle. When $\kappa > 1$ (bottom), there is weak spatial frequency overlap and the signal and pump break free from the respective potential wells created by their mutual interaction, but the signal still imparts some of its oppositely directed momentum to the pump such that the pump misses the spatial aperture. As shown in the figure, the signal does not completely escape the potential well of the pump and part of it remains bound to the pump. This is known as a daughter, or shadow, beam and can be removed by an analyzer at the output.

5.2 Soliton Collision and Dragging Logic Gates

This section evaluates in detail logic gates based on the phase-insensitive, inelastic interaction between orthogonally-polarized spatial solitons. In particular, the collision and dragging interactions are examined, because they belong to the class of angular deflection gates that were discussed in section 1.5.2, and shown to possess desirable properties. Note that the attraction and repulsion interactions are special cases of the more general collision interaction and will not be treated separately. The result of the comparison between the collision and dragging based logic gates in section 5.2.3 is that the dragging gates are generally shorter and are less sensitive to parameter variations.

For all simulations, as before, the normalized interaction parameters κ and r_s are used, where κ is the normalized angle defined by equation 5.4, and the initial soliton separation is given by $r_s w_0 [1 + 1/r]/2$. The length of a logic gate is denoted in terms of the signal's confocal distance, or $L = dZ_0$ where L is the total gate length. The pump soliton width is given by $w_p = w_0/r$ where r is the initial pump to signal soliton power ratio; therefore, if $P_0 = 2n_0 D / n_2^2 k_0^2 w_0$ is the initial power of the fundamental signal soliton where D is the waveguide thickness and n_2^2 is the Kerr index in intensity units, then $P_p = rP_0$ is the initial power of the fundamental pump soliton. The confocal distance of the pump is $Z_p = \pi^2 w_p^2 / \lambda = Z_0 / r^2$.

The previous section demonstrated that the inelastic interactions can produce sufficient angular deflection of the pump soliton to obtain large threshold contrast. Large threshold contrast provides an indication that small-signal gain (and therefore large-signal gain when $r > 1$) can be obtained. For a particular value of r , the gate transfer function is obtained as the plot of output power of the gate versus signal input power as illustrated in Figure 5.12, where the input signal amplitude is varied, but constant width is maintained such that the signal is a fundamental soliton at only one particular amplitude. At that particular amplitude, or power (to which the input and output levels are normalized), the threshold contrast τ is calculated. Therefore, for unity input level (given by P_0 , the power of the fundamental signal soliton), the normalized output level is given by $1/\tau$, and the transfer curve must pass through the point $(1, 1/\tau)$. When $\tau > 1$, only a fraction of the original signal soliton power (input level less than unity) is required to reach the threshold switching level, which is defined as the level at which the input and output are the same.

Since little deviation of the pump occurs for very small input level, the transfer curve is flat, or saturated, at the high output state. Therefore, the output remains saturated until an input level of P_{IL} is reached. Likewise, at high input levels beyond the point at which resolvable deviation of the pump occurs, further increase of the input leads to little change in the output. Here, P_{OL} is the output at this saturated level, and P_{IH} is the associated minimum input level. As a result, the transfer curve must also pass through (or near) the points $(0, P_{OH})$, (P_{IL}, P_{OH}) and (P_{IH}, P_{OL}) , where $P_{OH} = r$ in the absence of absorption. The heavy solid lines in Figure 5.12 indicate the canonical, or straight-line [109], representation of the transfer function that intersects the four known points, while the heavy dashed curve is representative of a more typical function. In terms of small-signal gain, the straight-line representation is the worst-case transfer function; any other transfer function must possess a region of larger small-signal gain.

For small-signal gain, the slope of the canonical transfer function, given by

$$\text{slope} = \frac{r - 1/\tau}{P_{IL} - 1}, \quad (5.21)$$

must satisfy the relation $|\text{slope}| > 1$, placing the following constraint on τ :

$$\tau > \frac{1}{r - |P_{IL} - 1|}. \quad (5.22)$$

When τ satisfies this criterion, a restoring logic gate (due to the presence of small-signal gain) should result for the given choice of parameters, assuming that the actual transfer function of the gate does not deviate too greatly from the piecewise approximation to the usual sigmoidal-type function as illustrated in the figure.

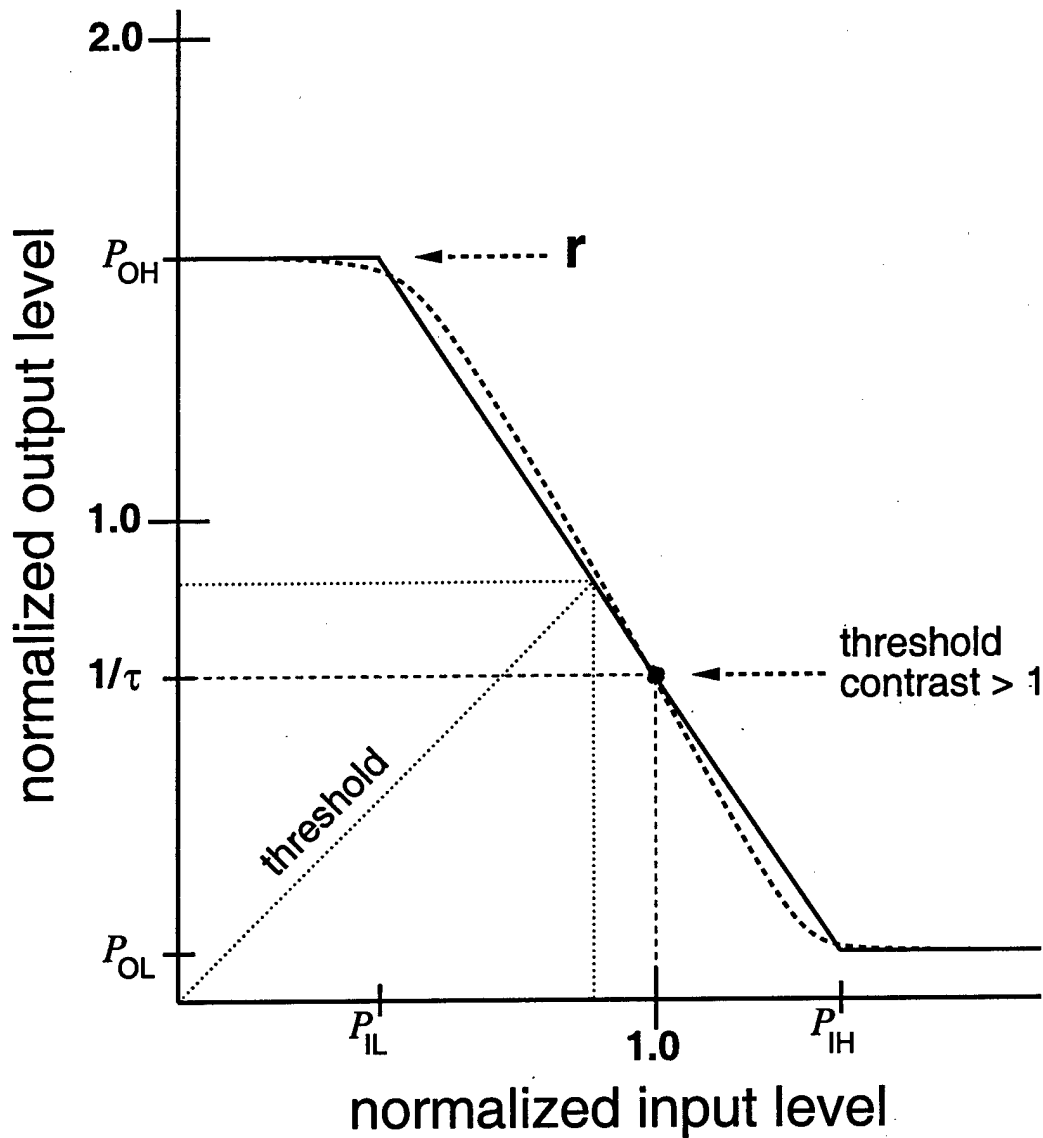


Figure 5.12: Generic gate transfer function. The threshold contrast τ is calculated at the unity input level (normalized by the fundamental signal soliton power), and $1/\tau$ gives the corresponding output level. The dotted lines indicate the switching threshold level where the input and output levels are equal. The heavy solid lines correspond to the worst-case (in terms of small-signal gain) transfer function, which must pass through the points $(0, P_{OH})$, (P_{IL}, P_{OH}) , $(1, 1/\tau)$ and (P_{IH}, P_{OL}) , while the dashed curve is representative of a more realistic transfer function, which necessarily has greater small-signal gain.

The final point to address is the choice of an acceptable value for τ . Based on equation 5.22, the minimum τ can be calculated when an estimate of P_{IL} is given. For the soliton dragging and collision logic gates studied here, P_{IL} can range from about 0.1 to about 1.0, with smaller values being more typical in high gain situations. The wide variation of the values of r and P_{IL} make the condition given by equation 5.22 difficult to generalize to a single value, or even range of values, of τ , but it should be clear that in all cases, larger values of τ are better. Another, more useful, measure is obtained by considering the situation when $1/\tau \leq P_{OL}$. Then it is clear that $P_{IH} \leq 1$. In this situation, an approximate lower bound to the large-signal gain is given by r . It is found in later simulations that typically $P_{OL} \sim 0.1$, so that $\tau > 10$ provides an indication that large-signal gain of at least r (derated by material absorption and the throughput factor of the output aperture) can be achieved. This does not mean that large-signal gain cannot be achieved when $\tau < 10$, because large-signal gain of at least unity is obtained as long as $P_{IH} < r$, which could occur for small τ . By equation 5.22, when $\tau > 10$, small-signal gain is obtained in every case for large r . Therefore, the condition $\tau > 10$ provides a strong indication that a gate will function with small- and large-signal gain, but it should be emphasized that $\tau < 10$ does not guarantee that a gate will not possess these properties.

The complete gate transfer function is only calculated in a few representative cases in section 5.4 where small-signal gain, large-signal gain, and gate fanout are formally defined, while evaluation of logic gates with more extensive variation of input parameters is based on the simple threshold contrast metric, which is sufficient for comparison purposes. In these situations, the minimum desired value of τ is chosen to be 10, as discussed previously.

5.2.1 The Soliton Collision Logic Gate

This section calculates the threshold contrast metric for the collision interaction for a range of gate lengths d , normalized interaction angles κ , and initial power ratios r . The operating regions are determined within which a restoring logic gate with small and large signal gain is expected.

A contour plot of threshold contrast τ versus gate length and initial power ratio is shown in Figure 5.13, indicating the regions within which threshold contrast exceeds a given value. Each sample point in the figure represents the maximum contrast obtained for an optimal value of κ . The first heavy contour represents the minimum desired contrast of 10; therefore, the region enclosed within this contour (which extends towards the upper right corner) meets or exceeds this criterion, indicating that large signal gain of about r should be obtainable as long as $P_{OL} \geq 0.1$. The contour levels are separated by powers of 10, so very large threshold contrasts are obtained for longer gate lengths and larger values of r . In these regions of very high contrast, for example within the second heavy contour representing $\tau \geq 10^4$, it is expected that gate operation is well within the saturated low output level and that large signal gains greater than r can be obtained. It is shown in section 5.4 that in these regimes of operation, $P_{OL} \sim 0.1$, such that $P_{IH} < 1$ because $1/\tau \ll P_{OL}$ at the unity input level.

The minimum gate length, defined as the length for which the threshold contrast equals 10, is a strong function of the initial power ratio r for the collision interaction. It is clear that when $r \leq 1.5$, the gate length increases rapidly with decreasing r , to the point that the minimum gate length is about $10 Z_0$ when $r = 1$. For the case when $r = 1$ at large angles, the colliding solitons tend to pass through each other in a similar manner to the elastic collision interaction between same-polarized solitons, as previously shown in the top of Figure 5.8. Therefore, the main result of the interaction is a small spatial shift in the propagation of the pump. At small angles, the solitons tend to form a bound orbiting pair as shown in the bottom of Figure 5.8, again with little induced change in the propagation angle of the pump. These behaviors manifest themselves in Figure 5.13 by noting that the contrast is a weak function of gate length when $r = 1$, a result of the small deflection angle of the pump. It should also be noted that unity contrast contour extends to zero gate length. This is because the choice of aperture size allows 94.2% of the undeviated pump power to pass. In the case of weak interaction then (i.e. $r = 1$), the lower bound on the threshold contrast is given by $\tau \geq 1/0.942 = 1.06$.

The minimum useful value of r for the collision interaction is about $r = 3$, for which the gate length decreases to about $5 Z_0$; a somewhat arbitrary criterion at this point, but justified later in section 5.2.3. When $r > 1$, the inelastic collision interaction tends to produce behavior that is analogous to the slingshot effect in orbital dynamics [104]. In this situation, the signal soliton enters the potential well of the pump (created through nonlinear cross-phase modulation, and a three-dimensional function of soliton widths and separation) at a shallow angle. The signal is not trapped in the potential well, and instead makes a partial orbit and emerges with oppositely directed transverse velocity. This dynamic is illustrated in the bottom interaction of Figure 5.9. It is also possible that the signal undergo multiple orbits before escaping [104], but only the simplest case is considered here. For a purely elastic interaction, twice the initial transverse momentum (derated by the cross-phase modulation coefficient) of the signal would be transferred to the pump, resulting in a large permanent angle change. In the case of inelastic interaction studied here, part of the signal may remain bound within the potential well of the pump, and some may even pass through the potential well completely without orbiting, as shown in the top of Figure 5.9, resulting in less than twice the transverse momentum of the signal being transferred to the pump. For fixed initial signal transverse momentum, a larger fraction of the signal remains bound to the pump with increasing r .

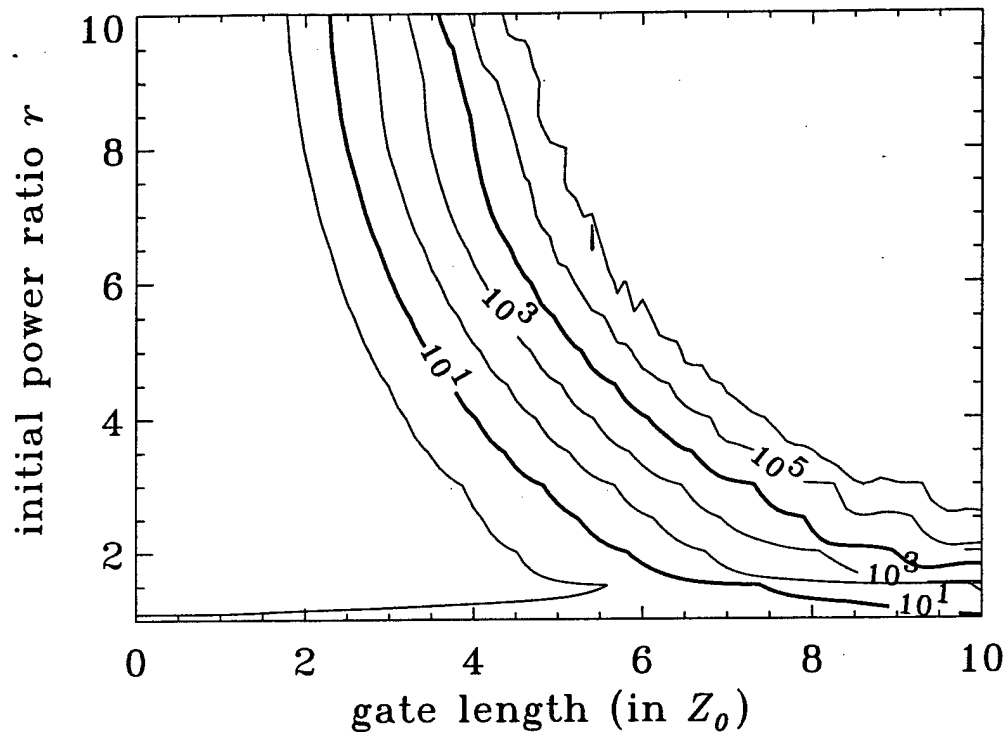


Figure 5.13: Threshold contrast versus gate length and initial power ratio for the collision interaction. The contours enclose regions where the maximum contrast (for individually optimal values of κ) exceeds the given value. The contours are spaced in decades and the thick contours represent threshold contrast levels of 10 and 10^4 . The minimum desirable contrast is 10, and the normalized separation parameter $r_s = 5$.

For the angular deflection logic gates, the threshold contrast will increase with increasing gate lengths, because a fixed angle change will result in greater spatial shift. It is not immediately obvious, however, that threshold contrast increases with r for a fixed gate length, as indicated in Figure 5.13. There are four factors that give rise to this behavior. These factors are linked by the fact that the pump width w_p decreases with r . At each value of r , the aperture width is maintained at $3.5w_p = 3.5w_0/r$, meaning that approximately a resolvable spatial shift is needed to obtain high contrast. The required spatial shift decreases with r because of the scaling of the aperture size. Therefore, a smaller absolute angle change (proportional to $1/r$ for small angles) is necessary to achieve that shift for a given fixed gate length. The second factor is the increase in the "mass" of the pump with r . This would indicate that for a fixed initial transverse momentum of the signal given by δk_x , the induced deflection angle of the pump decreases with r . These two factors suggest that the spatial shift (when normalized to the pump width) remains nearly constant with increasing r , but that the threshold contrast decreases with r because, for the same normalized spatial shift, more power will exit the aperture for larger r .

As shown in Figure 5.14 for two values of gate length, the optimal angle κ tends to remain localized about the average value κ_{avg} , for most values of r , resulting in an increasing transverse momentum, or real propagation angle, of the signal on average. Not all of this increased momentum is transferred to the pump, however, because part of the signal immediately escapes the potential well of the pump (due to the increase), some of the remaining signal performs a partial orbit and emerges in the opposite direction, and the rest remains bound. The result of nearly constant κ combined with increased breakup of the signal is that the actual deflection angle of the pump scales approximately as $1/r$, as expected, supporting the previous arguments. The fourth effect is that the normalized spatial separation is fixed at $r_s = 5$, meaning that the separation in physical units decreases as $[1 + 1/r]$. When combined with the increase in propagation angle of the signal, the actual interaction distance decreases, causing an effective increase in the gate length which leads to additional spatial shift. Because the pump is exponentially localized, only a slight additional spatial shift is needed to overcome the increase in transmitted power by the three previous arguments, and thus the decrease in effective interaction distance leads to the observed increase in contrast.

Another feature illustrated in Figure 5.14 is that the average value of κ for a gate length of $5 Z_0$ is larger than for a gate length of $10 Z_0$. The larger value of κ produces a slightly smaller deflection angle, leading to somewhat lower contrast for long

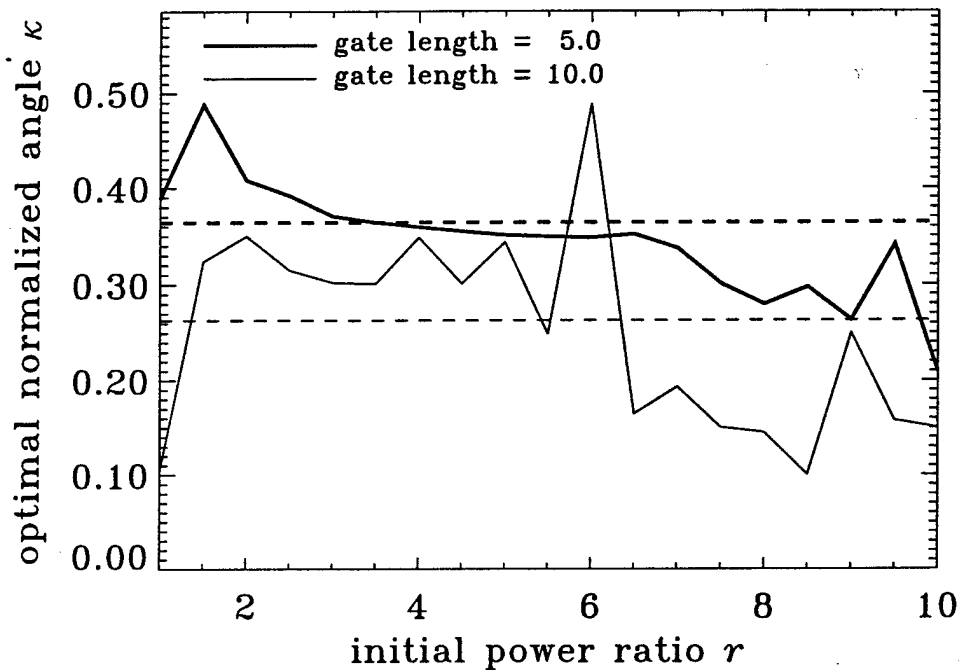


Figure 5.14: Plot of the optimal value of κ that produces the maximum threshold contrast, versus initial power ratio r for the collision interaction. Two curves are plotted, one (heavy linestyle) for a gate length of $5 Z_0$, and one (thin linestyle) for a gate length of $10 Z_0$. The dashed lines show the average value of κ for each gate length, with $\kappa_{\text{avg}} = 0.263$ for $10 Z_0$ gate length, and $\kappa_{\text{avg}} = 0.364$ for $5 Z_0$ gate length.

gate lengths, but decreases the effective interaction distance which leads to an increase in contrast for shorter gate lengths. A smaller value of κ results in increased interaction distance, but larger angular deflection compensates for the former effect to produce higher threshold contrast at the longer gate lengths.

5.2.2 The Soliton Dragging Logic Gate

This section calculates the threshold contrast metric for the spatial dragging interaction over the same range of gate lengths and initial power ratios as used in the previous section. The differences for the dragging interaction are that larger values of κ (~ 0.8 versus about 0.4) produce optimal results, and the minimum gate lengths are shorter.

Figure 5.15 shows the regions of initial power ratio and gate length in which the threshold contrast exceeds a given value. As before, the first heavy contour represents the minimum desired contrast of 10, and very large threshold contrast is obtained for longer gate lengths, as expected for an angular deflection gate.

Unlike the collision gate, the minimum gate length for the dragging interaction is relatively constant with initial power ratio. As discussed before in detail, this provides an indication that the deflection angle of the pump decreases approximately as $1/r$, which compensates for the scaling of aperture width, and results in a normalized (to the pump width) spatial shift that is nearly constant with respect to r . However, the spatial shift must increase slightly with r to compensate for the increased power of the pump. For complete trapping, as shown in the top of Figure 5.11, the induced angle change must in fact decrease with r in order that both solitons propagate at the weighted mean angle to conserve transverse momentum. However, as noted previously, the maximum contrast is obtained when spatial dragging occurs; that is, when all or part of the signal escapes the potential well of the pump and literally drags the pump to the side in the process, also inducing a permanent angle change that is not necessarily resolvable. This situation is shown at the bottom of Figure 5.11. Part of the signal, known as the soliton shadow, may also remain bound to the pump, resulting in an orbiting pair. This behavior is in part responsible for the numerous disjoint contour regions in Figure 5.15, but the major contributing factor is discussed in section 5.2.3.

As shown in Figure 5.16, the optimal angle κ tends to remain constant with r for the shorter gate length. Therefore, the initial transverse momentum of the signal increases with respect to r . Since the solitons initially overlap in space, the effective interaction distance (i.e. the distance over which the induced angle change becomes fixed) for a given value of r is determined

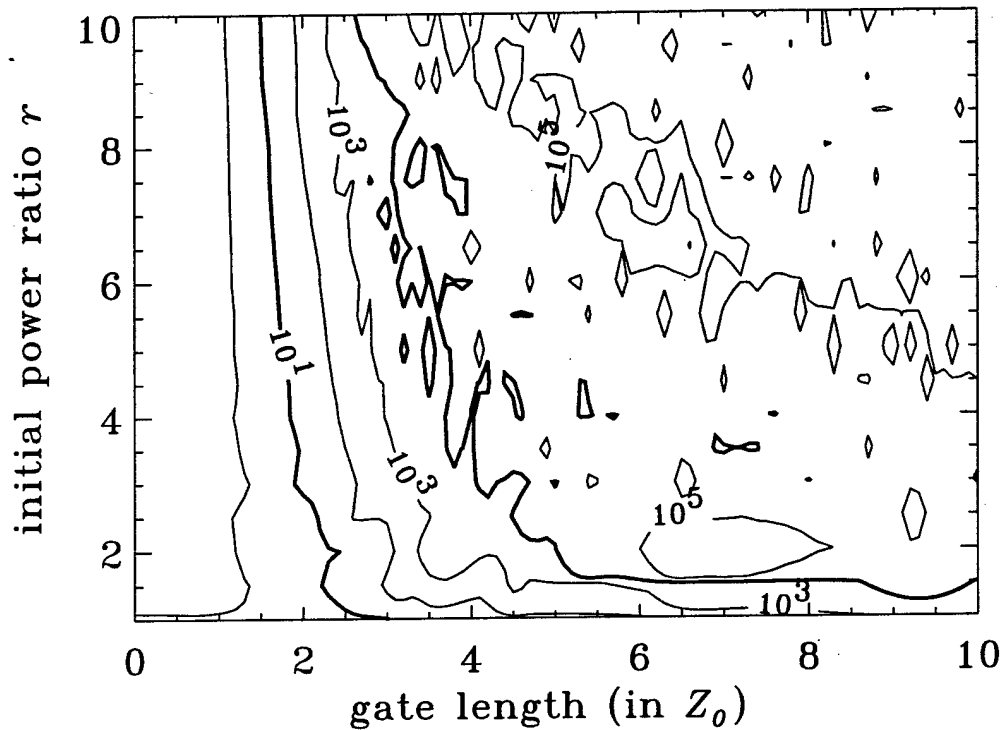


Figure 5.15: Threshold contrast versus gate length and initial power ratio for the spatial dragging interaction. The contours enclose regions where the maximum contrast (for individually optimal values of κ) exceeds the given value, and are spaced in decades. The thick contours represent threshold contrast levels of 10^4 and 10^5 . The minimum desirable contrast is 10, and the normalized separation parameter $r_s = 0$.

by the propagation angle of the signal, which does increase and cause a slight enhancement, but does not result in strongly enhanced contrast for increasing value of r due to simultaneous decrease in spatial separation as with the collision gate.

The large variation in the optimal value of κ for $10 Z_0$ gate length is due mainly to the orbiting between the pump and bound portion of the signal. The phase and period of this oscillation is a strong function of the initial conditions. The highest contrast is obtained when the pump is on the side of the signal which is opposite the aperture, and the optimal values of κ for a given gate length and ratio typically correspond to this condition. Therefore, there can be strong variation of κ with respect to gate length for fixed r and with respect to r for fixed gate length, especially at the longer gate lengths for which the oscillation is well developed.

5.2.3 Comparison of Logic Gates

This section provides a comparison between the collision and dragging interactions in terms of their properties for the implementation of logic gates. Particular attention is paid to the maximum value of r at a given gate length for which threshold contrast $\tau \geq 10$, and the minimum gate length for a given value of r . Both of these characteristics turn out to be important when the effects of material absorption are studied in section 5.3. Large values of r are generally needed to provide large fanout and to maximize the amount of pump throughput in the case of absorption, but optimized interaction parameters for small ratios can result in lower power dissipation in the presence of two-photon absorption and are shown in section 5.4 to provide sufficient fanout for most purposes. Small gate lengths are necessary to minimize gate latency and the effects of linear absorption.

Figure 5.17 is a contour plot of threshold contrast versus gate length and κ for the case when $r = 1$. Only results for the dragging interaction are shown, because the collision interaction with $r = 1$ cannot achieve the minimum contrast condition for the given parameters. The $\tau = 1$ contour is not present because when $r = 1$, the lower bound of $\tau \geq 1.06$ is placed due to the aperture.

The contour plot shows that the minimum length of the dragging gate is about $3 Z_0$ for unity initial ratio, and that the gate is insensitive to a wide range of variation in κ at longer gate lengths. The value $\kappa = 0.8$, as indicated by the dashed line, allows

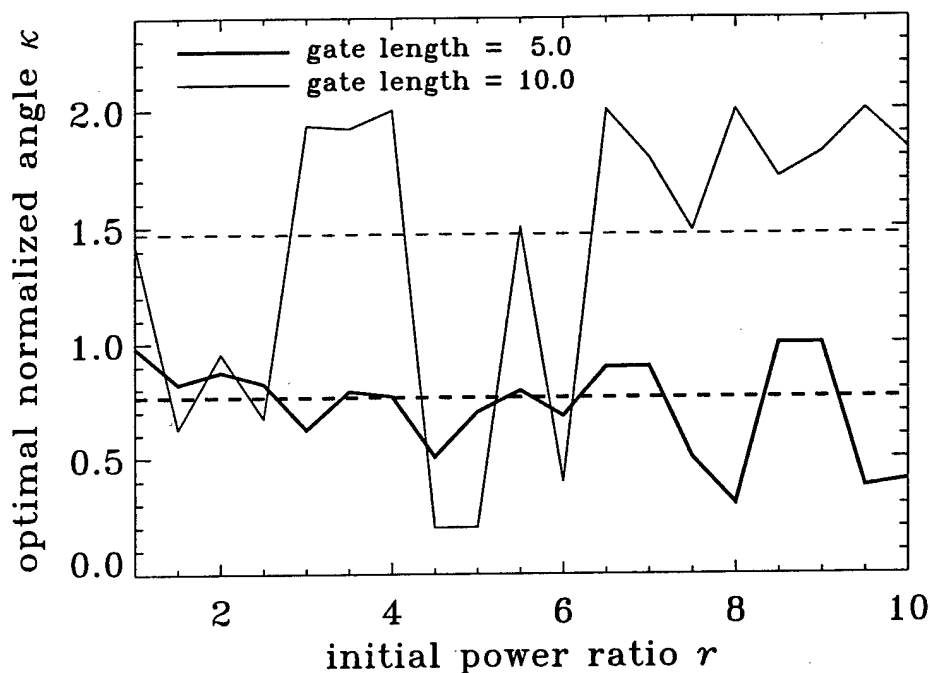


Figure 5.16: Plot of the optimal value of κ that produces the maximum threshold contrast, versus initial power ratio r for the spatial dragging interaction. Two curves are plotted, one (heavy linestyle) for a gate length of $5 Z_0$, and one (thin linestyle) for a gate length of $10 Z_0$. The dashed lines show the average value of κ for each gate length, with $\kappa_{\text{avg}} = 1.47$ for $10 Z_0$, and $\kappa_{\text{avg}} = 0.764$ for $5 Z_0$.

for high threshold contrast for all gate lengths greater than $3 Z_0$. This value will be used in section 5.4 to calculate the transfer function for this gate.

Figure 5.18 shows similar plots for both the collision and dragging gates when $r = 3$. The plots clearly indicate that shorter gate lengths can be obtained with the spatial dragging interaction. In fact, the minimum dragging gate length is about $2 Z_0$, while the minimum collision gate length is about $5 Z_0$. This large difference manifests itself in the performance of the gates in the presence of material absorption, as discussed in section 5.3.

The main reason why minimum soliton collision gate length is greater than the minimum dragging gate is that in the collision interaction the solitons are not overlapping at the input. As a result, the actual interaction occurs about $1-2 Z_0$ inside the material. This extra distance could be reduced by decreasing the initial separation at the input such that there is some, but not complete, overlap. However, the distinction between the collision and dragging interactions becomes blurred, such that they may be no longer considered separate. Perhaps the main point of this argument is to reveal the tolerance to (asymmetric) spatial misalignments that these deflection interactions possess.

Another important difference between the interactions is that the operating regimes of the dragging gate are much less sensitive to variations in the normalized angle κ than for the collision gate. As a result, the collision gate will require more precise tolerancing in a real implementation. The large acceptance of the dragging gate to κ also suggests another advantage over the collision gate in that multiple signal inputs can be used to implement a multi-input NOR gate in a single stage. In this implementation, the signals must be greater than linearly resolved (i.e. differences in κ greater than 1) in order to avoid phase sensitive behavior. Figure 5.18 shows that this indeed should be possible and that high contrast can be obtained when either signal is present (for example, $\kappa = 0.8$ and $\kappa = 2.0$), for gate lengths of about $4.5 Z_0$ or longer. Of course, the case when both signals are present needs to be investigated in order to guarantee robust operation, but this study is beyond the focus of this thesis.

Even in the inelastic collision interaction, the pump soliton tends to return to its original shape well after the collision. This means that almost no dispersive radiation is generated from the pump and that the threshold contrast increases exponentially for more than spatially resolvable deflection of the pump due to exponential localization. This explains the regular contours for the collision gate in Figures 5.13 and 5.18 and the fact that the maximum threshold contrast is more than an order of magnitude higher than for the dragging gate. The dragging interaction is inelastic for the pump soliton as well as for the signal. The result

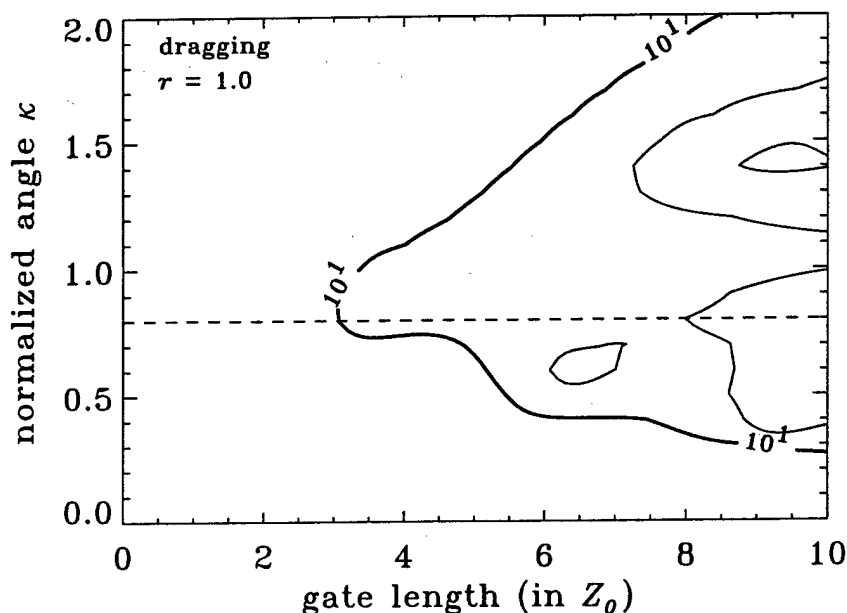


Figure 5.17: Threshold contrast contours for the spatial dragging interaction for $r = 1$. The contours are spaced by powers of 10 and the heavy contour represents the minimum desired contrast of 10. Only the $\tau = 10, 100$, and 1000 contours appear in this plot. The dashed line represents a value of κ of 0.8 that serves as a good operating point for all gate lengths.

is that during the interaction, dispersive radiation is generated from the pump, leading to exponential localization of the main radiation with well defined side lobes. These side lobes are typically three- to four-orders of magnitude smaller in intensity than the main pump, but in the large contrast regime $\tau > 10^3$ for large deflection, they form the major contribution to the power that flows through the aperture. This effect is seen in the contours of Figures 5.15 and 5.18. The difference between the maximum obtainable contrast for the collision and dragging gates is not a major factor because the gate fanout calculated in section 5.4 does not change noticeably between $\tau = 10^3$ and $\tau = 10^4$ for example.

Figure 5.19 shows example geometries for collision and dragging gates. In each case, $r = 3$ and the gate length is $10 Z_0$. The normalized interaction angles correspond to those indicated by the dashed lines in Figure 5.18. The plots clearly show that large deflection of the pump soliton occurs in both interactions, leading to high contrast operation.

The $r > 1$ collision interaction is highly inelastic for the signal. In the case shown here in the top plot, the width of the deflected pump is the same as that at the input, but at the point of collision, the signal breaks up into dispersive waves, with spatial broadening of the reflected signal wave clearly shown. Note that in some cases (not shown in Figure 5.19 for collision), part of the signal will remain bound to the pump as discussed previously, and result in the mutual focusing (termed here cross-focusing) of both the bound pump and signal. Since the reflected wave propagates in the opposite direction to the deflected pump, some power in the signal polarization may escape through the aperture. In order to block this leakage, an analyzer is required at the output, and is assumed present for all simulations.

The spatial dragging interaction is slightly inelastic for the pump as well, resulting in the side lobe generation discussed earlier. The bottom of Figure 5.19 shows that the signal soliton breaks up into two distinct waves. One dispersive wave propagates along the initial direction of the signal, while the other wave, the shadow, remains bound to the pump. This bound pair results in mutual cross-focusing, such that the final pump width decreases to $w_p = w_0/3.22$ (from $w_p = w_0/3$), and the final bound signal width decreases from w_0 to $w_0/2.90$. Since the signal always propagates to the same side of the aperture as the pump, no analyzer is required at the output to block leakage of the signal polarization.

5.3 Effects of Material Absorption

Strongly nonlinear optical materials invariably suffer from the presence of absorption. Resonant enhancement of the non-linearity directly leads to an increase in the linear absorption and an increase in the response time (due to the decay of real particle excitation in the dynamic nonlinear regime); both effects are detrimental to ultrafast all-optical switching. It is well

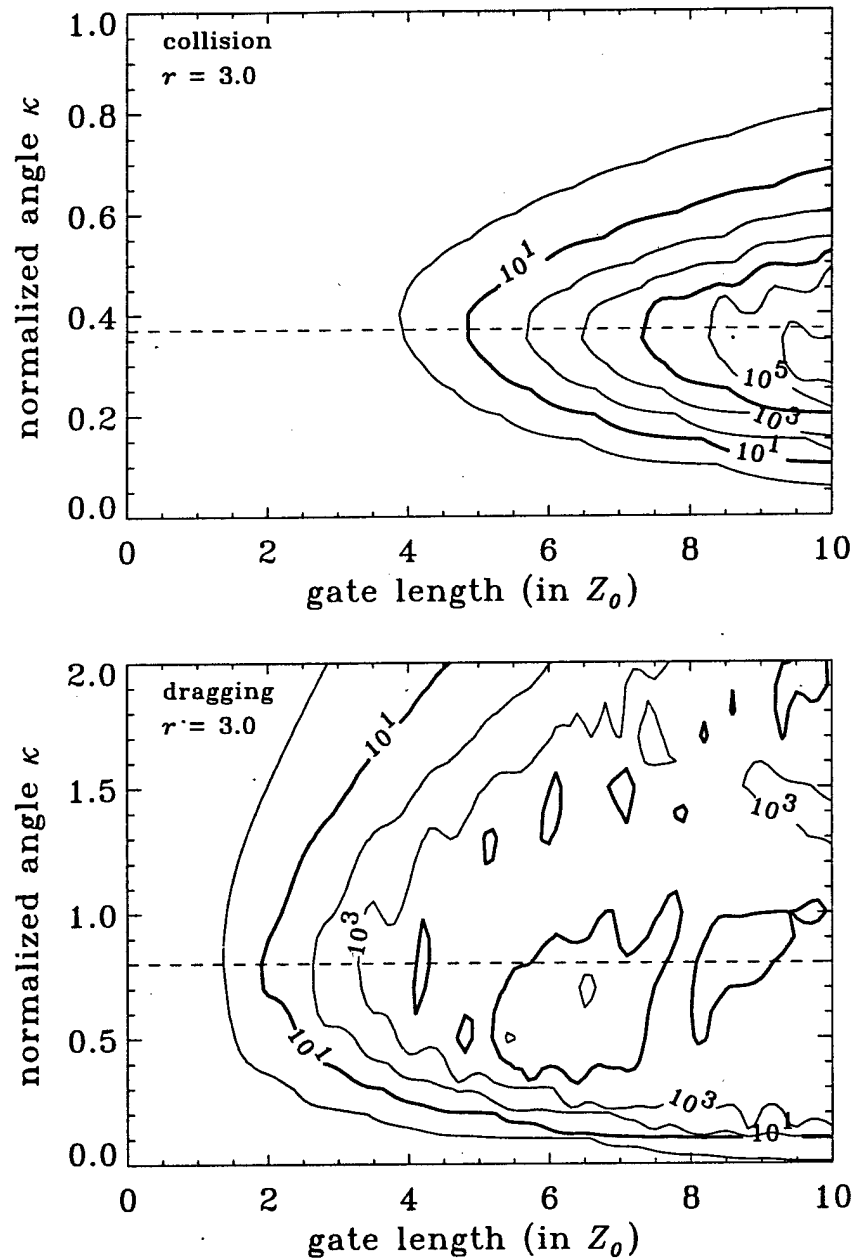


Figure 5.18: Threshold contrast contours for the spatial collision (top) and dragging (bottom) interactions for $r = 3$. The contours are spaced by powers of 10 and the heavy contours represents the minimum desired contrast of 10, and a contrast of 10^4 . The dashed lines represent values of κ which serve as good operating points for all gate lengths. In the case of the collision interaction, this value is 0.37, and in the case of the dragging interaction, this value is 0.8.

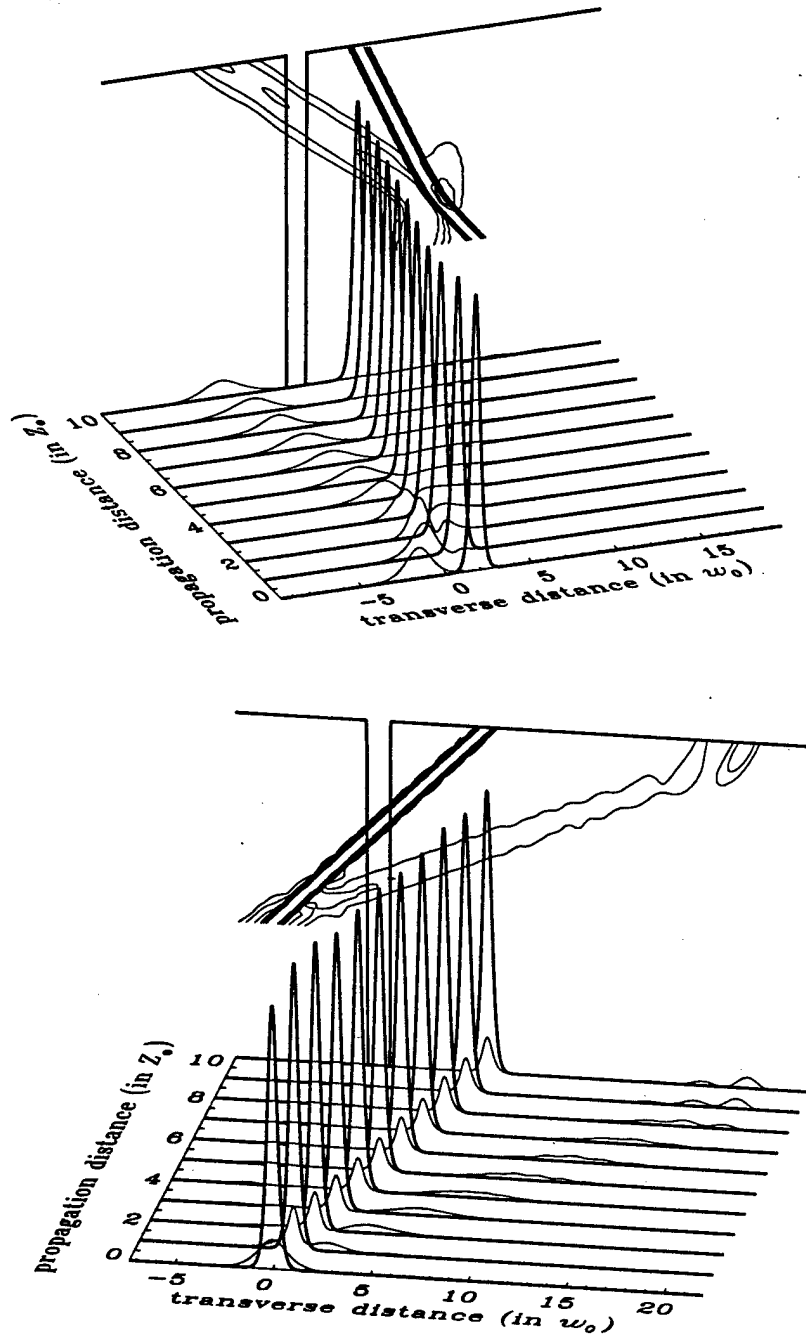


Figure 5.19: Spatial soliton collision (top) and dragging (bottom) logic gates. Here, the gate lengths are $10 Z_0$, the initial ratio $r = 3$, and the normalized initial separation is $r_s = 5$ for collision and $r_s = 0$ for dragging. The normalized interaction angle is $\kappa = 0.37$ for collision and $\kappa = 0.8$ for dragging. The pump soliton is represented by the heavy linestyle and the contours are spaced in 3 dB intervals. The threshold contrast of the collision gate is 1.6×10^6 , and 1.4×10^4 for the dragging gate.

known that linear absorption places fundamental restrictions on the operation of nonlinear devices such as bistable Fabry-Perot etalons [110] and directional couplers [111]. These studies have produced a nonlinear Fabry-Perot figure of merit $3n_2^2 I / \alpha_0 \lambda \gg 1$ [110] where n_2 is the nonlinear Kerr index, α_0 is the linear absorption constant, $\lambda = \lambda_f / n_0$ is the material wavelength and I is the intensity, and a nonlinear directional coupler material figure of merit $\Delta\beta_{\text{sat}} / \alpha_0 \lambda_f \gg 1$ [112], where $\Delta\beta_{\text{sat}}$ is the maximum nonlinear change in the effective propagation constant and λ_f is the free-space wavelength. In order to minimize linear absorption, a promising regime of operation in nonlinear optics is near half-bandgap in semiconductors [113] or in the transparency regions of glasses and organics. These non-resonant nonlinearities are essentially instantaneous with very small linear absorption due mainly to material or waveguide scattering, but with some non-zero two-photon absorption. The effects of two-photon absorption on nonlinear directional couplers [114] have been studied and a switching criterion has been defined as $T = 2\beta_2 \lambda_f / n_2 \ll 1$ [115, 116], where β_2 is the two-photon absorption coefficient.

Bright soliton propagation in the presence of linear absorption was first studied to determine the implications for temporal soliton transmission in fiber [117]. More recent work has focused on the effects of nonlinear absorption on spatial soliton propagation. The effects of two-photon absorption on bright spatial soliton propagation were observed experimentally in glass waveguide experiments [118], and then studied analytically [119]. Beam propagation has been used to successfully model soliton breakup due to two-photon absorption [119, 120]. Stationary propagation of dark solitons was predicted in the presence of both two-photon absorption and linear amplification [121], and the effects of two-photon absorption on soliton nonlinear interface switches [122] and dark soliton steering [123] have also been studied. In this section, the separate and combined effects of linear and two-photon absorption on bright spatial soliton propagation and the implications for spatial soliton logic gates are discussed in detail. Figures of merit are defined in order to evaluate the suitability of a given material for use in logic applications, and the operation of spatial soliton collision and dragging logic gates in the presence of absorption using realistic material parameters is demonstrated, again using threshold contrast as the performance metric.

5.3.1 Adiabatic Evolution of a Single Soliton

The development starts by considering the effects of absorption on the propagation of a single soliton. This analysis results in simple relationships that place upper bounds on the ratio of pump power at the output aperture to the fundamental input signal soliton power, and therefore places limitations on the large signal gain. The effects of absorption on the spatial collision and dragging interactions will be studied numerically.

The (1+1)-D spatial nonlinear Schrödinger equation with the addition of linear and two-photon absorption terms is

$$2ik_0 \frac{\partial A}{\partial z} + \frac{\partial^2 A}{\partial x^2} + 2k_0^2 \frac{n_2}{n_0} [1 + iK] |A|^2 A + ik_0 \alpha_0 A = 0 \quad (5.23)$$

where k_0 is the material propagation constant, $K = \beta_2 / 2k_f n_2$ [119], β_2 is the two-photon absorption coefficient, n_2 is the nonlinear Kerr index, and α_0 is the linear absorption coefficient.

The equation of motion for the 1-D integrated power of the nonlinear wave, with the boundary conditions $A(x = \pm\infty) \rightarrow 0$, is

$$\frac{\partial}{\partial z} P(z) \equiv \frac{1}{2\eta} \frac{\partial}{\partial z} \int_{-\infty}^{\infty} |A|^2 dx = -\frac{\alpha_0}{2\eta} \int_{-\infty}^{\infty} |A|^2 dx - \frac{\beta_2}{2\eta} \int_{-\infty}^{\infty} |A|^4 dx, \quad (5.24)$$

where the relation $\beta_2 = 2k_0 K n_2 / n_0$ was used and $\eta = 1/\epsilon_0 c n_0$ is the material impedance (in ohms, neglecting the nonlinear contribution to refractive index) which converts V^2/m^2 to units of W/m^2 . Equation 5.24 shows that the decrease in integrated power with respect to propagation is linearly proportional to the power by the constant α_0 and nonlinearly proportional to the two-photon absorption constant β_2 . In the absence of absorption, the change in integrated power is zero and equation 5.24 corresponds to the conservation of energy.

An adiabatic evolution assumption is made such that the fundamental soliton maintains the characteristic sech() shape during propagation. Therefore, in order to include the effects of absorption on soliton propagation, the following ansatz is used

$$|A(x, z)| = \frac{1}{k_0 w(z)} \sqrt{\frac{n_0}{n_2}} \text{sech} \left[\frac{x}{w(z)} \right] \quad (5.25)$$

where the width parameter $w(z)$ is the only soliton parameter and is assumed to vary adiabatically with propagation, $w_0 \equiv w(0)$ is the initial fundamental soliton width, and the initial power $P_0 = 1/2\eta \int |A|^2 dx = 2n_0/k_0^2 n_2^2 w_0$. Substituting equation 5.25 for the fundamental soliton into the equation of motion 5.24, the following expressions for the soliton width and power are obtained

$$w^2(z) = \frac{4K}{3\alpha_0 k_0} [e^{2\alpha_0 z} - 1] + w_0^2 e^{2\alpha_0 z} \quad (5.26)$$

$$P^2(z) = \frac{P_0^2}{k_0^3 [n_2^I]^2 K P_0^2 [e^{2\alpha_0 z} - 1] / 3\alpha_0 n_0^2 + e^{2\alpha_0 z}}. \quad (5.27)$$

Before considering the general case, the two special cases of equations 5.26 and 5.27 of linear and two-photon absorption in isolation are considered.

Linear Absorption Only

The first case is when $\beta_2 = 0$, corresponding to negligible two-photon absorption. Then the evolution equation for the soliton width and power become

$$w(z) = w_0 e^{\alpha_0 z} \quad P(z) = P_0 e^{-\alpha_0 z} \quad (5.28)$$

where the power attenuation is the same as for linear plane waves as expected. For a propagation distance of $z = dZ_0$ and attenuation over that distance of e^{-s} , the relationship $s = \alpha_0 dZ_0$ provides a constraint on the initial soliton width

$$w_0 \leq \sqrt{\frac{s\lambda}{\pi^2 \alpha_0 d}} \quad (5.29)$$

where the maximum value of s depends on the amount of linear absorption that an interaction can tolerate. For a fixed, allowable value of s , large linear absorption α_0 must be compensated by a small width w_0 (to a minimum of about λ) to reduce the propagation distance such that the linear attenuation stays bounded by e^{-s} .

The soliton width and power can be re-written

$$w_0(s, d) = \mathcal{F}_\alpha \lambda \quad P_0(s, d) = \frac{\lambda_f}{2\pi^2 n_2^I \mathcal{F}_\alpha} \quad (5.30)$$

in terms of the linear absorption figure of merit \mathcal{F}_α , which is defined as

$$\mathcal{F}_\alpha(s, d) \equiv \sqrt{\frac{s}{\pi^2 \alpha_0 d \lambda}}. \quad (5.31)$$

The restriction $\mathcal{F}_\alpha > 1$ is placed in order that the soliton width parameter w_0 be greater than about a material wavelength. When $w_0 = \lambda$ for example, the linear angular divergence of the fundamental soliton $\text{sech}()$ eigenmode is $\approx \pm 10^\circ$, so the paraxial approximation is at the edge of validity; clearly, for $\mathcal{F}_\alpha < 1$, the paraxial approximation begins to break down, and the additional non-paraxial terms derived in Chapter 3 need to be included in the nonlinear wave equation. The non-paraxial linear terms are included in all subsequent simulations, as discussed in section 4.2.1.

Since the pump is a fundamental soliton of narrower width than the signal (i.e. $w_p = w_0/r = \mathcal{F}_\alpha \lambda/r$), the maximum initial ratio of the pump to the signal is given by $r = \mathcal{F}_\alpha$, so that $w_p > \lambda$. Therefore, an upper bound to the ratio of undeviated pump power at the output to the initial fundamental signal soliton power at the input is given by $r(d) = \mathcal{F}_\alpha e^{-s}$. When the threshold contrast for a gate $\tau \geq 10$, then $r(d)$ (which is $\propto \sqrt{s}e^{-s}$) is the approximate lower bound to the large-signal gain, or fanout. Clearly, this lower bound needs to be maximized, which occurs when $s = 0.5$, resulting in $r(d) = 0.1365/\sqrt{d\alpha_0\lambda}$.

Figure 5.20 shows the propagation over $10 Z_0$ of a fundamental signal, an $r = 3$ fundamental pump, and an $r = 4$ higher-order pump (for later comparison with propagation in a two-photon absorbing medium) in a linearly-absorbing material where $s = 0.5$. Notice that the "soliton" width increases smoothly for the initial fundamental eigenmodes, such that the ansatz 5.25 is valid.

Two-Photon Absorption Only

The other special case is when the material has negligible linear absorption, but high enough intensities are used such that two-photon absorption becomes significant. Then $\alpha_0 = 0$ and the expressions for the width and power reduce to

$$w^2(z) = w_0^2 + \frac{8Kz}{3k_0} \quad (5.32)$$

$$P^2(z) = \frac{P_0^2}{2k_0^3 [n_2^I]^2 K P_0^2 z / 3n_0^2 + 1}, \quad (5.33)$$

which were first obtained in normalized form [119]. After a propagation distance of dZ_0 , the power of the fundamental signal can be written

$$P^2(d) = \frac{P_0^2}{dT/6 + 1} \quad (5.34)$$

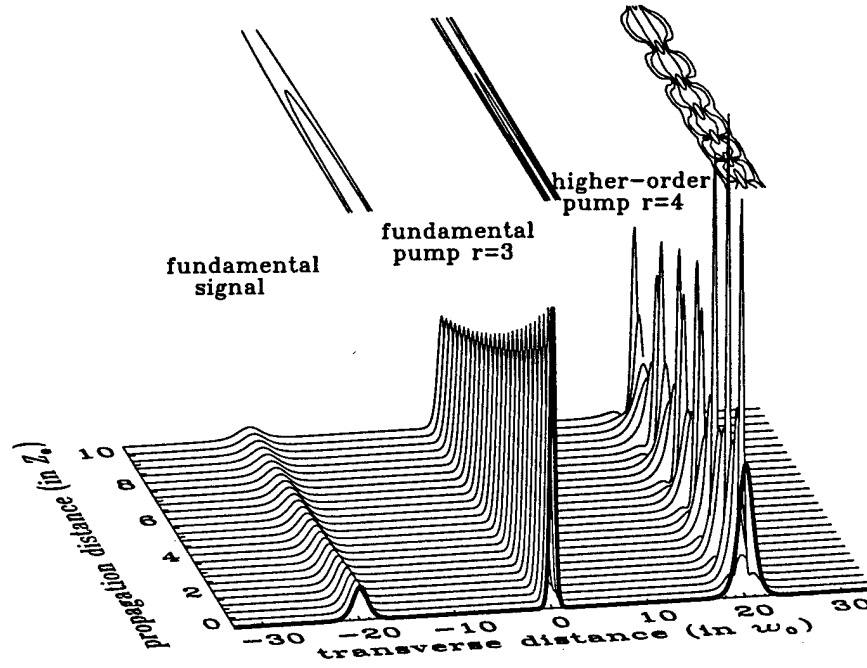


Figure 5.20: Propagation of a fundamental signal soliton, an $r = 3$ fundamental and $r = 4$ higher-order pump solitons in a linearly-absorbing material with $s = 0.5$. The propagation distance is $10 Z_0$. With linear absorption, the signal and pumps attenuate by the same fraction of $e^{-0.5} = 0.61$.

where $T = 2\beta_2\lambda_f/n_2 = 8\pi K$ is the two-photon absorption material figure of merit for a nonlinear directional coupler [116]. The attenuation of the signal is independent of its width w_0 , and therefore initial intensity, which is a result of the confocal distance (and absolute gate length) scaling as the square of the width and the peak intensity scaling as the inverse square.

In the presence of two-photon absorption, higher-order solitons have been shown to break up during propagation [118–120], thereby not providing a cascable high output state for a logic gate. This is one reason why a fundamental soliton of narrower width than the signal is used for the pump in order to achieve $r > 1$, where the integrated pump power is given by $P_p = rP_0$. The power of the pump after propagating the gate length dZ_0 is then

$$P_p^2(d) = \frac{r^2 P_0^2}{(dT/6)r^2 + 1}. \quad (5.35)$$

The maximum pump power (as $r \rightarrow \infty$) at the output of the gate (just before the aperture) can be written as

$$P_p(d)_{\max} = \sqrt{\frac{6}{dT}} P_0. \quad (5.36)$$

The two-photon absorption figure of merit for soliton interactions is defined as the maximum ratio of the pump at the output $P_p(d)_{\max}$ to the signal at the input P_0 , or

$$r(d) = \mathcal{F}_{\text{TPA}}(d) \equiv \frac{P_p(d)_{\max}}{P_0} = \sqrt{\frac{6}{dT}}. \quad (5.37)$$

In the absence of linear absorption, \mathcal{F}_{TPA} provides an approximate lower bound on the gate fanout that an interaction can provide when $\tau \geq 10$, and is a result of the optical limiting nature of two-photon absorption.

Equation 5.35 for the power of the pump at a propagation distance dZ_0 can be rewritten in terms of the figure of merit and the initial power ratio r as

$$P_p(d) = \frac{r \mathcal{F}_{\text{TPA}}(d) P_0}{\sqrt{r^2 + \mathcal{F}_{\text{TPA}}^2(d)}}. \quad (5.38)$$

Figure 5.21 shows a plot of this relationship versus d and parameterized by r for a material where $\mathcal{F}_{\text{TPA}}(10) = 10.8$, which corresponds to fused silica. It is clear that the majority of the absorption for high pump to signal beam ratios takes place at

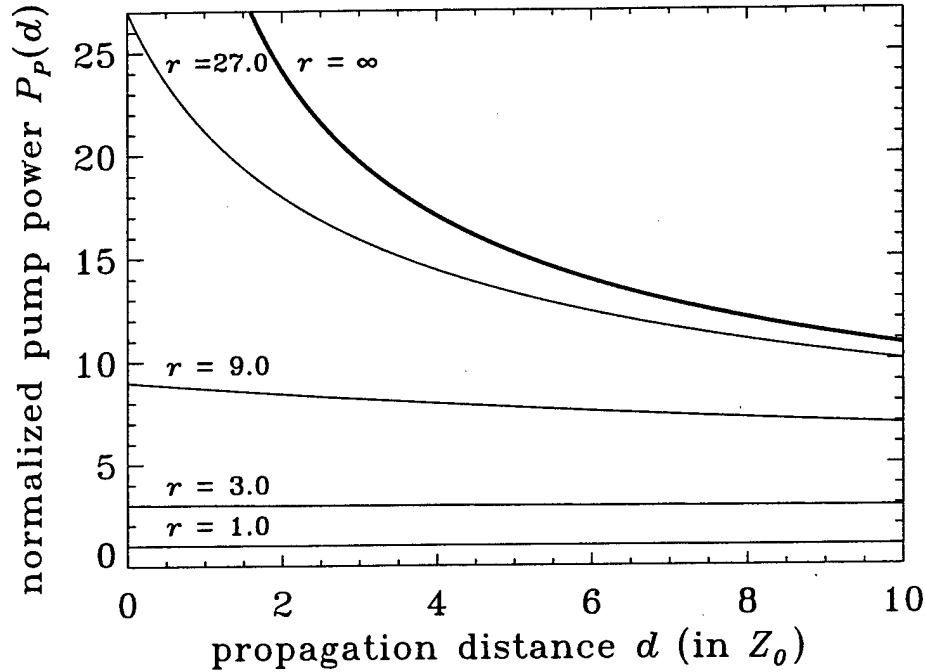


Figure 5.21: Plot of pump power $P_p(d)$ given by either equation 5.35 for $T = 0.00517$ or equation 5.38 for $\mathcal{F}_{\text{TPA}}(d) = 34.1/\sqrt{d}$ versus normalized propagation distance d (where $z = dZ_0$). The thin solid lines are the evolution of the pump power for a given initial beam power ratio and the thick solid line is the maximum (as $r \rightarrow \infty$) given by $\mathcal{F}_{\text{TPA}}(d)$.

the beginning of propagation, so that longer propagation distances do not pose as much of an additional penalty as with linear absorption, for which the absorption rate is constant. It is also clear that for large r , the power of the pump at the output asymptotically approaches the maximum given by $\mathcal{F}_{\text{TPA}}(d)$. For $\mathcal{F}_{\text{TPA}} \gg 1$, the attenuation at the signal level ($r = 1$) becomes negligible.

Figure 5.22 shows the propagation over $15 Z_0$ of a fundamental signal, a narrower $r = 3$ fundamental pump, and an $r = 4$ higher-order pump with the same width as the signal, for $K = 0.005$ (or $T = 0.13$). The signal power at some normalized gate length d follows equation 5.34, therefore the attenuation of the signal is negligible (as shown in the figure) when $dT < 6$, as is the case here for $d = 15$ and $dT = 2.0$. The fundamental pump is absorbed rapidly over the first $5 Z_0$ with much less attenuation over the next $10 Z_0$. The width of both fundamental beams increases smoothly with propagation. The longer propagation distance was chosen so that the splitting of the $r = 4$ higher-order pump would be clearly shown. At $z = 10 Z_0$, for example, the higher-order pump soliton reforms into a narrow peak that could be passed through the output aperture, but slight fluctuations in the input pump power would alter the distances over which coalescing occurs, thus leading to blocking of the split pump even in the absence of the signal. It is clear from this discussion and the figure that a higher-order pump cannot be used for restoring, cascaded logic in a material with non-negligible two-photon absorption.

Two-Photon Absorption and Plane Waves

The nonlinear directional coupler figure of merit T was derived with a plane-wave analysis ignoring the effect of attenuation on the accumulated nonlinearly-induced phase shift [116]. The power (in some finite area A_{eff}) of a plane wave due to two-photon absorption is given by

$$P_{\text{plane}}(z) = \frac{P_0}{1 + \beta_2^I I_0 z} \quad (5.39)$$

where $I_0 = P_0/A_{\text{eff}}$. Using the plane-wave model and thus ignoring any transverse variation, efficient switching in a nonlinear directional coupler requires that $T < 0.7$, fully accounting for the decreasing rate of accumulated nonlinear phase due to absorption throughout the device. Bright spatial soliton propagation in the presence of two-photon absorption should include

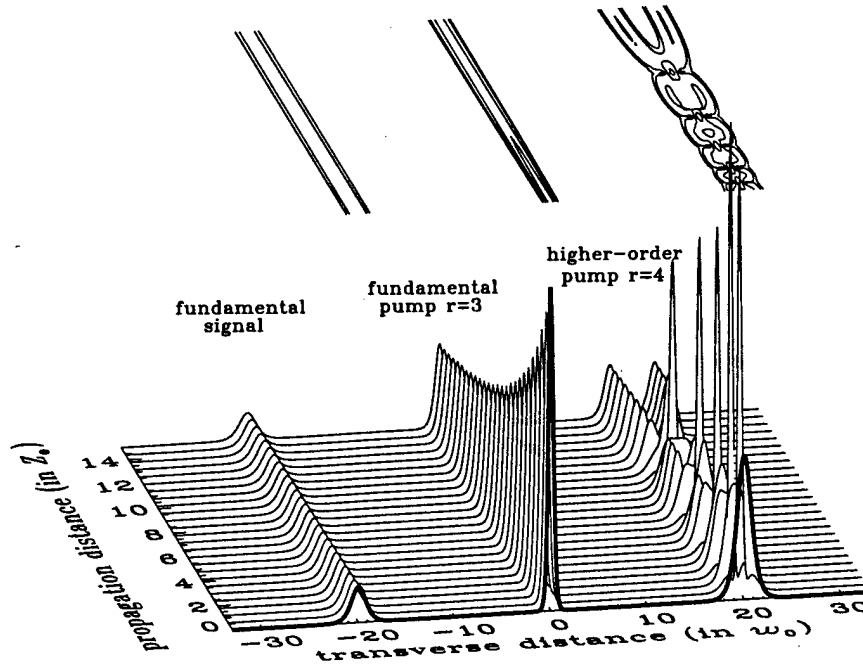


Figure 5.22: Propagation of a signal, an $r = 3$ fundamental pump and an $r = 4$ higher-order pump in a two-photon-absorbing material with $K = 0.005$ over a distance of $15 Z_0$. The longer distance is used to illustrate the higher-order soliton breakup more clearly. With two-photon absorption, the signal and pumps attenuate by different fractions; the signal by 0.87, the $r = 3$ fundamental pump by 0.51 and the $r = 4$ higher-order pump by 0.62.

the transverse variation of the field which accounts for the square-root relationship to the T parameter shown in equation 5.37. For a 1-D soliton, the peak intensity decreases as a result of two effects: the attenuation due to two-photon absorption and the adiabatic broadening of the beam due to the tendency to maintain the characteristic soliton shape. The smaller the peak intensity, the less attenuation, therefore, as a result of the beam broadening, the power attenuation for the same initial peak intensity I_0 over a given distance is then less for a soliton than for a plane wave as evident from a comparison of equations 5.33 and 5.39. Note that equation 5.33 can be rewritten as

$$P_{\text{soliton}}(z) = \frac{P_0}{\sqrt{1 + (4/3)\beta_2^I I_0 z}}, \quad (5.40)$$

which asymptotically approaches zero with increasing distance much more slowly than the plane-wave propagation given by equation 5.39, thus making soliton-based devices more tolerant to two-photon absorption in terms of power loss than non-soliton devices.

Simultaneous Linear and Two-Photon Absorption

In many cases, even operating in a transparency region of a material, there is some linear absorption due to material or waveguide scattering (typically $< 1 \text{ cm}^{-1}$). Considering both linear and two-photon absorption, the power of the pump at some distance dZ_0 is given by

$$P_p(s, d) = \frac{r\sqrt{2s} \mathcal{F}_{\text{TPA}}(d) P_0}{\sqrt{r^2 [e^{2s} - 1] + 2s \mathcal{F}_{\text{TPA}}^2(d) e^{2s}}} \quad (5.41)$$

which reduces to equation 5.38 for $s \ll 1$. Now the initial pump to signal beam ratio r is limited to \mathcal{F}_α , so that

$$r(d) = \mathcal{G}(s, d) = \frac{\sqrt{2s} \mathcal{F}_\alpha(s, d) \mathcal{F}_{\text{TPA}}(d)}{\sqrt{\mathcal{F}_\alpha^2(s, d) [e^{2s} - 1] + 2s \mathcal{F}_{\text{TPA}}^2(d) e^{2s}}}, \quad (5.42)$$

Material	fused silica	39% Pb silicate	PTS	Al ₁₈ Ga ₈₂ As
s_{opt}	0.263	0.0679	0.132	0.5
$\mathcal{F}_{TPA}(5)$	15.2	4.15	4.06	$\gg 1$
$\mathcal{F}_\alpha(s_{opt}, 5)$	22.3	15.2	10.0	38.5
$\mathcal{G}(s_{opt}, 5)$	10.5	3.86	3.48	23.3 ^a
$\mathcal{F}_{TPA}(10)$	10.8	2.94	2.87	$\gg 1$
$\mathcal{F}_\alpha(s_{opt}, 10)$	15.7	10.7	7.08	27.2
$\mathcal{G}(s_{opt}, 10)$	7.42	2.73	2.46	16.5 ^a

^a given by $\mathcal{F}_\alpha e^{-s}$

Table 5.1: Calculated figures of merit for promising nonlinear materials for gate lengths of $5 Z_0$ and $10 Z_0$.

which reduces to $\mathcal{G}(s, d) = \mathcal{F}_{TPA}(d) \sqrt{2s/[e^{2s} - 1]}$ for large \mathcal{F}_α , and approaches \mathcal{F}_{TPA} for small s .

The ratio $r(d)$ given by equation 5.42 can be maximized by the value of s that satisfies

$$\frac{e^{-2s_{opt}}}{1 - 2s_{opt}} = \left(\frac{12\pi^2 \alpha_0 \lambda}{T} + 1 \right), \quad (5.43)$$

which depends only on material parameters. Notice that $0 \leq s_{opt} < 0.5$ for $T > 0$, and $s_{opt} = 0.5$ when $T = 0$ as previously shown with linear absorption in isolation.

5.3.2 Figures-of-Merit for Example Materials

This section evaluates the linear and two-photon absorption figures of merit for a few promising materials for use in soliton logic gates (and many nonlinear devices in general). The relevant linear and nonlinear properties of these materials are given in Table D.1. The primary example used for the studies of this thesis is fused silica, mainly because of its large transparency region in the visible and near infrared, the presence of anomalous dispersion for use in the spatio-temporal case, and for the ease with which devices can be fabricated. Another study focused on lead-doped silicate glass [70], which has smaller figures of merit.

For fused silica at $\lambda_f = 1.55 \mu\text{m}$ [30], the linear index $n_0 = 1.444$, the nonlinear Kerr index $n_2^I = 3.3 \times 10^{-16} \text{ cm}^2/\text{W}$, the two-photon absorption coefficient $\beta_2^I = 5.5 \times 10^{-15} \text{ cm/W}$, and it is assumed that the linear absorption coefficient $\alpha_0 = 0.1 \text{ cm}^{-1}$ due to scattering in a slab waveguide [124]. These material constants give $T = 0.00517$ and $\mathcal{F}_{TPA}(10) = 10.8$ as introduced previously. The upper bound on the ratio of the undeflected pump at the output to the signal at the input, or $r(d)$, is then determined by the value of the linear absorption constant s chosen, and hence the linear absorption figure of merit \mathcal{F}_α . The maximum value of s that is tolerable is determined by the interaction (note that the interaction degradation due to two-photon absorption must also be considered), but the optimal value of s in terms of maximizing the gain upper bound given by equation 5.43 is determined only by the material parameters. For the fused silica example, Figure 5.23 plots equation 5.31 for \mathcal{F}_α and equation 5.42 for the gain upper bound, versus s and parameterized by two normalized propagation distances, $d = 5$ and 10 . For both device lengths, the optimal value is $s = 0.2627$ as given by equation 5.43, providing an output pump to input signal ratio of $r(5) = 10.5$ when $r = \mathcal{F}_\alpha(0.2627, 5) = 22.3$, and $r(10) = 7.42$ when $r = \mathcal{F}_\alpha(0.2627, 10) = 15.7$.

Table 5.3.2 contains calculated figures of merit, where \mathcal{F}_α and \mathcal{G} are obtained for s_{opt} , for the four example nonlinear materials. Since the gain upper bounds are greater than unity, any of these four materials should be able to support multi-level, cascaded spatial soliton-based switching. Even though the glasses have large figures of merit, such high powers may be required to create a single soliton due the small nonlinear coefficients that the implementation of practical, low power systems may be impossible. As a result, it appears that PTS could be more useful because of its similarly large figures of merit and three-orders-of-magnitude larger nonresonant nonlinearity [125].

Operating below half-bandgap in semiconductors is promising for nonlinear optics because it is expected that two-photon absorption is zero in this regime. Due to defect states, though, there can be the equivalent of an Urbach tail in the two-photon absorption that extends below half-bandgap [126]. The AlGaAs data is taken for a sample of sufficient purity in which two-photon absorption is for all practical purposes zero [127]. In this case where linear absorption is dominant, the optimal value $s = 0.5$ results in $\mathcal{F}_\alpha(0.5, 10) = 27$ with $r(10) = 17$. The large linear absorption figure of merit indicates that there is a lot of headroom to reduce s and still maintain $\mathcal{F}_\alpha(s, 10) > 1$, but the soliton power and peak intensity increase with decreasing \mathcal{F}_α . Since two-photon absorption is nearly zero, the effects of three-photon absorption may need to be considered. To estimate the magnitude of this effect, the three-photon absorption figure of merit [128] $V = \lambda_f \beta_3^I I_0 / n_2^I = n_0^2 \beta_3^I / 2\pi k_0 [n_2^I]^2 \mathcal{F}_\alpha^2$ is calculated,

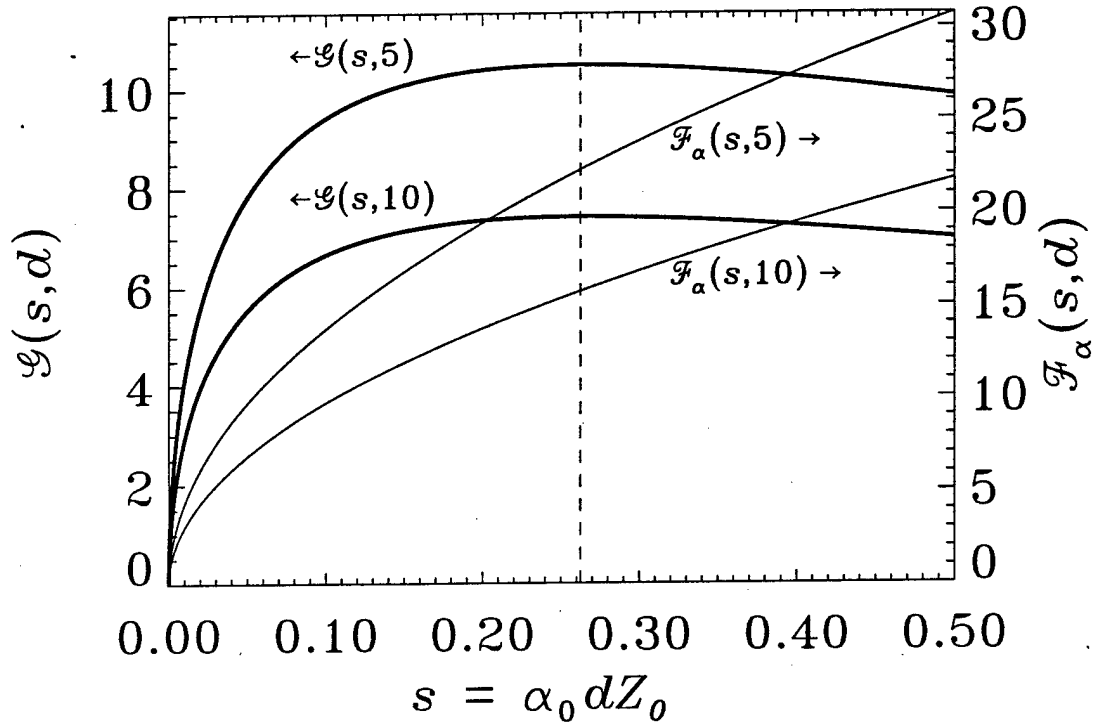


Figure 5.23: Plot of the linear-absorption figure of merit \mathcal{F}_α (thin line) given by equation 5.31 and gain upper bound \mathcal{G} (heavy line) given by equation 5.42 versus the linear absorption parameter s parameterized by normalized gate lengths $d = 5$ and $d = 10$. The material parameters for fused silica are used and the optimum gain upper bound occurs in both cases at $s_{opt} = 0.2627$, as given by equation 5.43. For these parameters, $\mathcal{G}(0.2627, 5) = 10.5$ and $\mathcal{G}(0.2627, 10) = 7.42$.

and should be much smaller than unity. At $d = 10$, $\mathcal{F}_\alpha(0.5, 10) = 27$ and $V = 0.07$ for $\beta_3^I = 0.055 \times 10^{-18} \text{ cm}^3/\text{W}$ [127], indicating that three-photon absorption can be neglected for the present study. Lowering s and hence lowering \mathcal{F}_α will increase V .

The material examples presented thus far do not make use of resonant enhancement. The material restrictions placed by the linear absorption figure of merit have implications to resonantly enhanced nonlinearity. At the very minimum, a logic gate needs unity fanout (actually a fanout of 2 is required for the synthesis of arbitrary logic function with a complete gate family). This requirement taken with the expression for the gain upper bound in a linearly-absorbing material $\mathcal{F}_\alpha(0.5, d)e^{-0.5} = 0.1365/\sqrt{d\alpha_0\lambda}$, places the restriction on the linear absorption constant $\alpha_0 \leq 0.0186/d\lambda$. As will be shown in the next section, the approximate minimum normalized gate length for the dragging interaction is $d = 5$ when $s = 0.5$. It is also not unreasonable to assume a material wavelength of $\lambda = \lambda_f/n_0 = 0.3 \mu\text{m}$. This means that to obtain a soliton dragging logic gate with $r(5)$ at least unity requires a material with linear absorption constant $\alpha_0 \leq 125 \text{ cm}^{-1}$ and for $r(5) = 2$, $\alpha_0 \leq 31 \text{ cm}^{-1}$. This clearly indicates the importance of experimental determination of the linear absorption spectrum and nonlinear dispersion curve in order to achieve the maximum nonlinearity within the absorption constraint. Thus the tradeoff between low absorption and large nonlinearity will have to be made in favor of tuning far out on the wings of the one-photon resonant absorption peak.

5.3.3 Comparison of Logic Gates with Absorption

Now, the operation of the spatial soliton collision and dragging gates is studied in the presence of absorption. The phase-insensitive interaction between orthogonally-polarized solitons is modeled by the coupled nonlinear Schrödinger equations

$$2ik_0 \frac{\partial A_x}{\partial z} + \frac{\partial^2 A_x}{\partial x^2} + 2k_0^2 \frac{n_2}{n_0} [1 + iK] \left[|A_x|^2 + 2\Delta |A_y|^2 \right] A_x + ik_0 \alpha_0 A_x = 0 \quad (5.44a)$$

$$2ik_0 \frac{\partial A_y}{\partial z} + \frac{\partial^2 A_y}{\partial x^2} + 2k_0^2 \frac{n_2}{n_0} [1 + iK] \left[|A_y|^2 + 2\Delta |A_x|^2 \right] A_y + ik_0 \alpha_0 A_y = 0 \quad (5.44b)$$

where any material and/or form birefringence has been neglected along with the phase-dependent term in the nonlinear polarization. Even though large birefringence is necessary to strongly phase-mismatch the vector four-wave mixing terms, the birefringence terms can be neglected as well because the evolution equations are only coupled through the optical intensity. Here, $2\Delta = 2/3$ for an isotropic material and orthogonal linear polarizations away from 1- and 2-photon resonances.

The two-photon absorption for the x -polarization consists of two parts: the simultaneous absorption of two x -polarized photons and the cross-absorption of one x -polarized and one y -polarized photon. Since the real and imaginary parts of $\chi^{(3)}$ are subject to the same intrinsic symmetries, Δ is the same for both cross-phase modulation and cross-two-photon absorption. This additional term in the nonlinear absorption is illustrated by writing the analog of equation 5.24 for the total integrated power of each polarization of the coupled system

$$\frac{\partial}{\partial z} P_x(z) = -\frac{\alpha_0}{2\eta} \int_{-\infty}^{\infty} |A_x|^2 dx - \frac{\beta_2}{2\eta} \int_{-\infty}^{\infty} (|A_x|^4 + 2\Delta |A_y|^2 |A_x|^2) dx, \quad (5.45)$$

where P_x is the power contained in the x -polarization and the coupled equation for P_y is similar. The cross-two-photon absorption of a probe beam interacting with a stronger pump was used recently [129] to measure the two-photon absorption anisotropy in cubic GaAs and CdTe. The same effect occurs here in the isotropic case in that the pump (signal) will absorb more strongly in the presence of the signal (pump) than without, but because $2\Delta = 2/3$ for an isotropic material when using orthogonal linear polarizations (sufficiently far away from two-photon resonance), the induced absorption caused by an orthogonally-polarized beam of the same intensity is slightly weaker than the self-induced absorption.

The same interaction parameters r_s and κ used before are employed here as well. The one difference is that only a single value for κ is used for each interaction. For collision, this value is $\kappa = 0.37$, and for dragging, $\kappa = 0.8$; these values are obtained from Figure 5.18, but are representative as well for $r \neq 3$. As before, the output aperture width is based on the pump soliton width, but now, it is the pump soliton width at the aperture, or $w_p(d)$, which is different than the width at the input due to absorption. The aperture width is set to $3.5w_p(d)$, where the pump width is calculated from equation 5.26, with $w_p = w_0/r$ substituted for w_0 in the expression. Because the pump can now propagate in the non-paraxial regime (because w_0 is fixed at a particular value determined by the material absorption and gate length), the simulations are performed using the equations described in section 4.2.1.

Using the material parameters of fused silica, contour plots of threshold contrast versus initial power ratio r and gate length are shown in Figure 5.24. For each gate length d in the figures, the fundamental signal soliton width w_0 is chosen such that the amount of linear absorption over the length of the gate is given by $e^{-s_{opt}}$, using the value $s_{opt} = 0.2627$. This initial signal width is calculated from the expression

$$w_0 = \sqrt{\frac{s_{opt}\lambda}{\pi^2\alpha_0 d}} = \mathcal{F}_\alpha \lambda. \quad (5.46)$$

This forces the maximum initial power ratio to \mathcal{F}_α for that gate length. Then, using expression 5.42, the maximum final power ratio $r(d)$ is given by $\mathcal{G}(s_{opt}, d)$, and indicated in the plots by the dashed contours. For small values of r , two-photon absorption is weak and the dashed contours are nearly horizontal.

Within the threshold contrast contour levels $\tau \geq 10$, $\mathcal{G}(s_{opt}, d)$ (multiplied by the power throughput factor of the aperture of 0.942) represents an approximate lower bound to the large-signal gain, or fanout, for that gate. Note that for the parameters of fused silica, the minimum value of $r = \mathcal{F}_\alpha(s_{opt}, d)$ is 15.7 when $d = 10$, which is larger than the highest initial ratio of 10 used in Figure 5.24. The main reason why the maximum value is not used is that sufficiently high values of $r(d)$, for which the interactions produce $\tau \geq 10$, can be obtained for large fanout, as shown in section 5.4. In fact, for the longer gate lengths, greater values for r do not produce significantly greater values for $r(d)$, this is the result of the rapid attenuation at high peak intensities due to two-photon absorption.

By comparing Figure 5.24 with Figure 5.13 for collision without absorption and with Figure 5.15 for dragging without absorption, it is seen that the main effect of absorption on both interactions is the reduction of threshold contrast. Although direct quantitative comparison is difficult because the previous figures showed contrast for optimal values of κ , while the present figure uses a single value of κ for each interaction, this simple observation deserves further discussion. Because of absorption of the undeflected pump at the output, the threshold contrast metric is automatically inflated by at least the factor $e^{s_{opt}} = 1.3$ (due to linear absorption alone). This factor is offset by the broadening of the widths of the interacting solitons. The signal soliton (and the pump soliton for small r) is not significantly affected by two-photon absorption and the broadening is due mostly to linear absorption alone, with the broadening factor given by $e^{s_{opt}}$. In the regime of linear absorption dominated broadening then, the final (undeflected) width of the pump is constant with respect to gate length, and a fixed angular deviation will result in greater spatial shift for longer lengths, as expected for the angular deflection gates. It should be pointed out, however, that because of the broadening of the pump, the normalized spatial shift of the pump for a given deflection angle and gate length is less than in the absence of absorption.

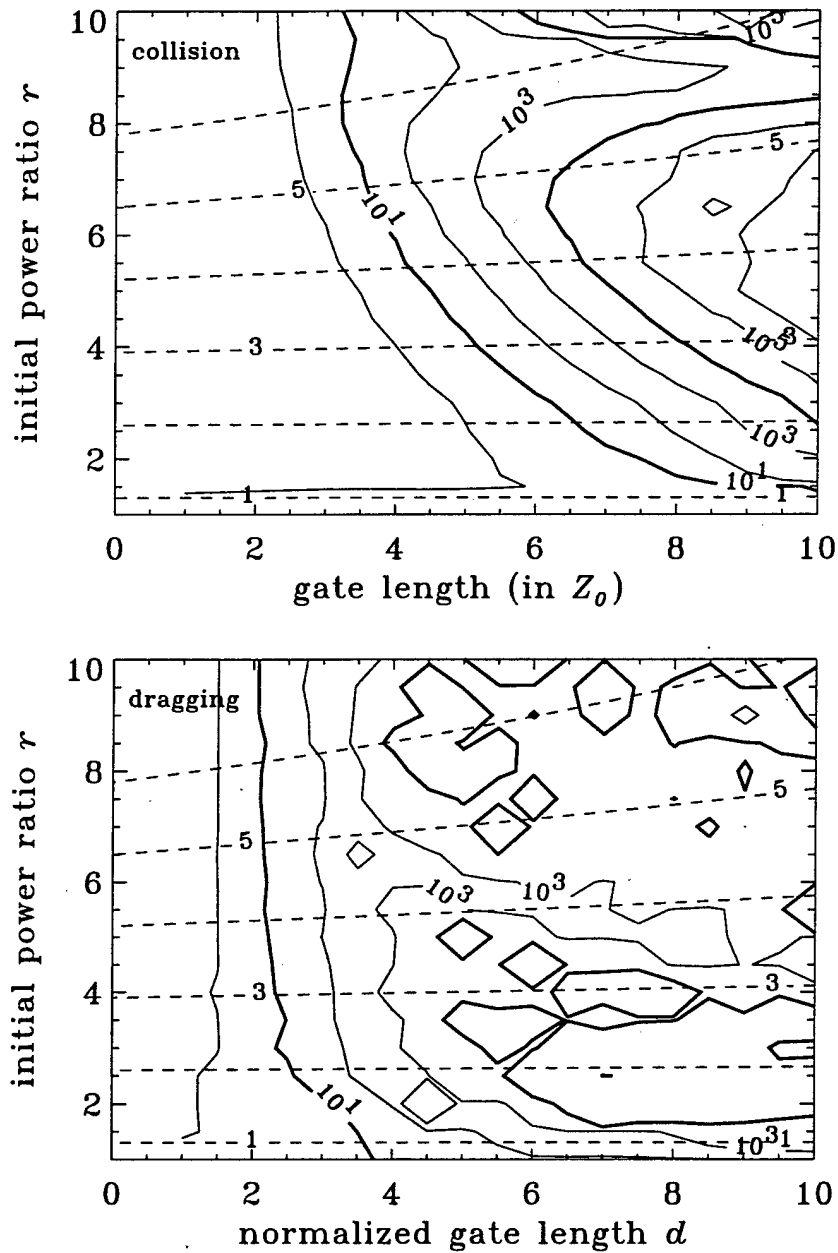


Figure 5.24: Threshold contrast contours for the collision (top) and dragging (bottom) gates using the parameters of fused silica and the value $s_{opt} = 0.2627$ for all gate lengths. The contour levels are spaced in decades and the $\tau = 10$ and $\tau = 10^4$ contours use the heavy linestyle. The normalized interaction angle is fixed at $\kappa = 0.37$ for collision and at $\kappa = 0.8$ for dragging. Absorption decreases the ratio of the undeflected pump at the output to the signal at the input. This ratio $r(d)$ is indicated by the dashed contours, which do not include the 5.8% power loss due to clipping by the output aperture.

For larger initial ratios (i.e. $r > 3$), the broadening of the pump is due to the combined action of linear and two-photon absorption, as given by equation 5.26. As a result of the pump width increasing with gate length due to two-photon absorption, the lever arm nature of the angular deflection gates is diminished because of the need for an increasing deflection angle with increasing gate length. The ramifications of this effect are seen in Figure 5.24 for the larger initial ratios and longer gate lengths. In these regions, the threshold contrast is significantly reduced over the case without absorption, and the effect is more pronounced for the collision interaction.

Like the case with no absorption, Figure 5.24 clearly shows that the dragging interaction produces shorter gate lengths, which is more apparent for small r . Absorption does not influence dragging so strongly (besides the general factors discussed in the previous paragraphs) because the interaction occurs at the beginning of the nonlinear material due to complete overlap, before significant absorption occurs. Collision, on the other hand, takes place within the material such that absorption has already attenuated the pump and signal, causing their widths to increase. This increase in width actually decreases the interaction distance somewhat, but the interaction will subsequently require a longer gate length to achieve high contrast after the interaction due to the scaling of the confocal distance with soliton width.

Figure 5.25 shows soliton collision and dragging gates of length $10 Z_0$ with $s_{opt} = 0.2627$ for fused silica, and initial pump to signal soliton power ratio $r = 3$. At the output of the gate, $r(10) = 2.24$, neglecting the aperture. The plots show that the pump is more strongly absorbed at the beginning of the interactions due to two-photon absorption. In the dragging interaction, cross two-photon absorption between the pump and signal shadow results in stronger attenuation of the signal compared to that which would occur with the signal alone, and in slightly more attenuation of the pump than occurs in the collision interaction, for which there is nominal trapping.

The main disadvantage to the use of the collision interaction is the longer gate lengths required to achieve the minimum threshold contrast criterion for a given value of r , or more appropriately in the case of absorption, $r(d)$. Shorter gate length leads to reduced gate latency, so the spatial dragging based gates have a decided advantage in this regard. Another factor is that stronger pumps (i.e. larger r) are required at the input for the collision gates, and therefore more total input power per switch, leading to more power dissipated as heat. This effect is negligible in the linear absorption dominated regimes, but becomes significant for large values of r and τ , and becomes much more pronounced in materials with smaller two-photon absorption figures of merit [70].

5.4 Cascading of Collision and Dragging Gates

The previous two sections studied two instantiations of the angular deflection soliton logic gates based on the spatial collision and dragging interactions. Using the notion of threshold contrast τ , it was argued that a restoring logic gate with small-signal gain and fanout of r can be obtained for $\tau \geq 10$. The soliton collision and dragging logic gates were shown to possess large regions in parameter space where this condition holds. The final points to address for the soliton logic gates are the calculation of the gate transfer function, which formally defines small-signal gain, large-signal gain, and noise margin, and demonstration of cascability and fanout in the ring-oscillator and multi-input NOR configurations, which serve as fundamental building blocks for more complex systems.

For the optical logic gates studied in this thesis, the parameter of ultimate interest is the fanout factor, which can only be defined in a cascaded system of logic gates. The maximum allowable fanout is restricted by the minimum allowable noise margins. The regions of valid high and low output levels, with the corresponding valid low and high input levels, as illustrated in Figure 5.26, are defined as the regions of the transfer curve for which the slope is less than unity, such that any variation about the stable operating points of the system (which must lie within these valid regions) is attenuated. Therefore, as long as the noise margins about the operating points are greater than the statistical variation in the input levels, the gates will function properly in a complex system. If no additional sources of noise were present in the system, because the local slope in the region near the stable input levels is less than unity, the variance in the fluctuation about the input levels would asymptotically collapse to zero as the digital signal propagates through a chain of gates. The input levels corresponding to the operating points of the device are denoted I_L and I_H , and locate the two self-consistent, stable points of the transfer function. The threshold level is also a self-consistent point of the transfer function, but it is not a stable operating point because the local slope, or small-signal gain, is greater than unity.

The noise margins of a particular device are traditionally defined for an electronic gate as [26, 109]

$$NM_L \geq P_{IL} - P_{OL} \quad (5.47a)$$

$$NM_H \geq P_{OH} - P_{IH}, \quad (5.47b)$$

where powers are used here as the relevant gate input and output quantities (for optics) instead of voltages, and NM_L and NM_H are the high and low level noise margins. For the present studies, these relations provide lower bounds on the noise margins,

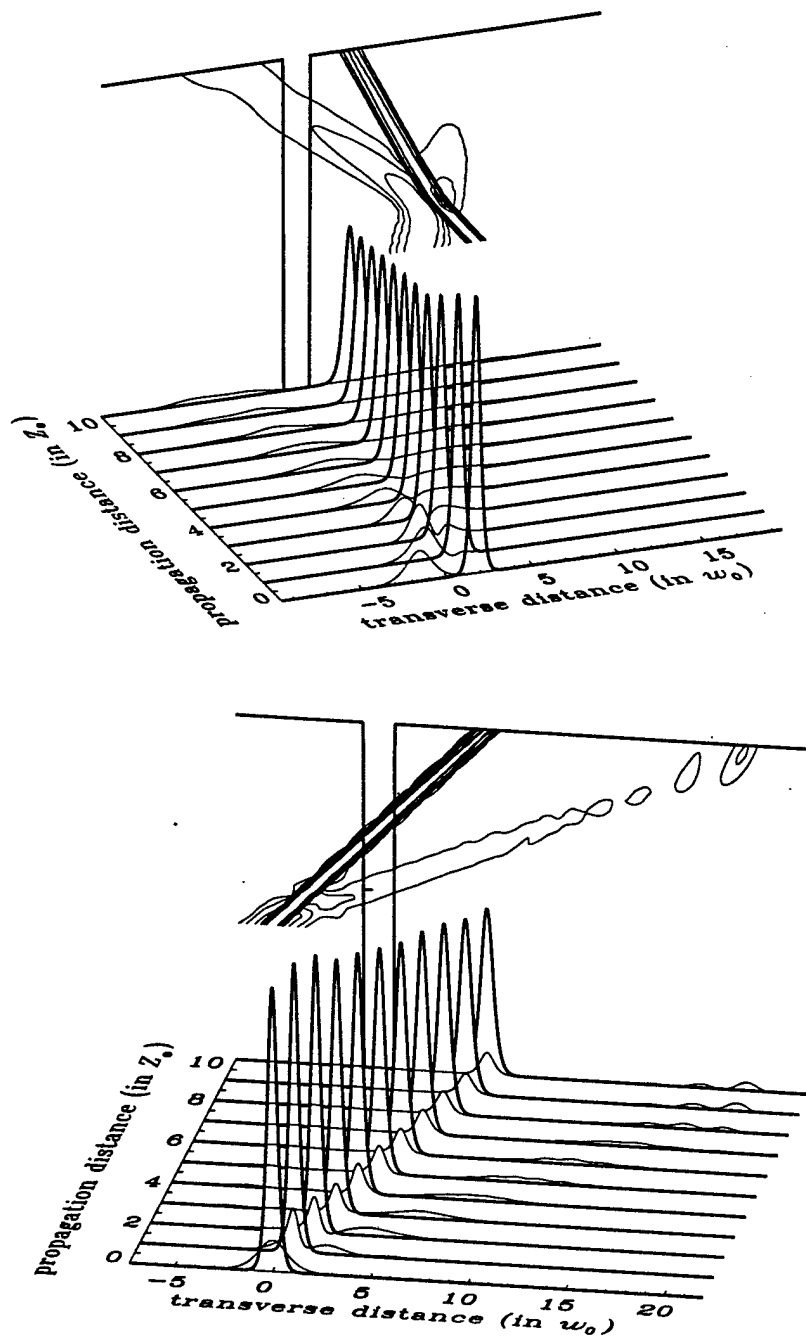


Figure 5.25: Spatial soliton collision (top) and dragging (bottom) logic gates in the presence of linear and two-photon material absorption. Here, the gate lengths are $10 Z_0$ with $s_{opt} = 0.2627$, the initial ratio $r = 3$, and the normalized initial separation is $r_s = 5$ for collision and $r_s = 0$ for dragging. The normalized interaction angle is $\kappa = 0.37$ for collision and $\kappa = 0.8$ for dragging. The pump soliton is represented by the heavy linestyle and the contours are spaced in 3 dB intervals. The threshold contrast of the collision gate is 3.1×10^4 , and 7.4×10^3 for the dragging gate.

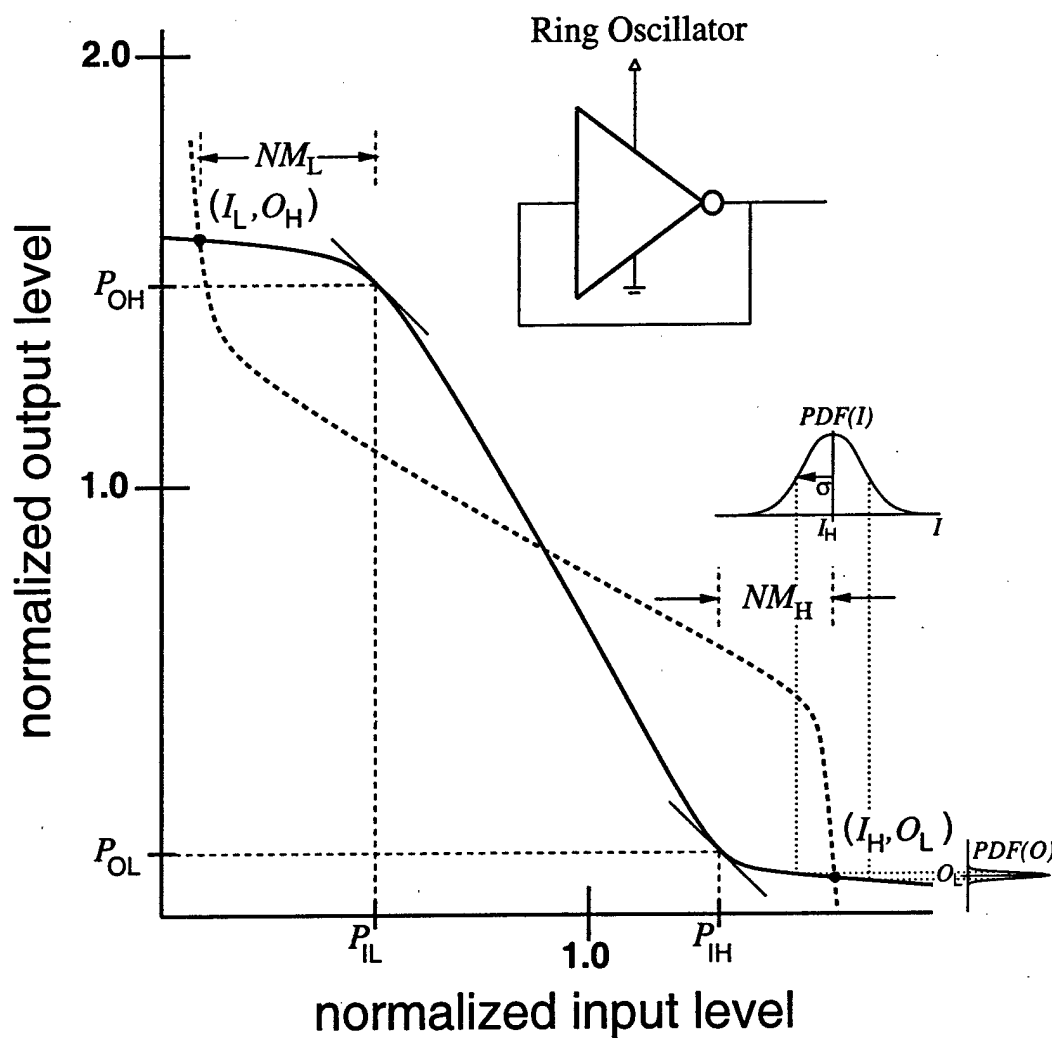


Figure 5.26: Gate transfer function illustrating the regions of valid input and output levels, which are defined as those regions on the transfer curve where the slope is less than unity. In the absence of fanout, the intersections of the normal and inverted transfer functions locates the stable operating points. The noise margins based on the operating points are also indicated, along with an example of the probability density function (PDF) representing the distribution of input levels about the operating point I_H .

which are more precisely defined based using the stable operating points, as illustrated in Figure 5.26, and discussed later. The probability density function representing the distribution of input signal values about the stable input levels I_L and I_H (which includes effects such as the laser relative intensity noise, scattering, variation in device parameters, etc) must fit within the noise margins of the gate to within a number of standard deviations σ . For instance, for a thousand gates, the noise margins about the stable operating points of the device must be at least 6σ ($\pm 3\sigma$ about the mean) wide [130]. For a particular value of fanout, the noise margins are re-calculated. The maximum fanout of the gate is obtained when the noise margins reach the minimum acceptable width in terms of σ . The statistical distributions about I_L and I_H are not considered in this thesis. Instead, based on the obtained transfer functions, the allowable noise margins are calculated, and it is argued that sufficient noise margin exists for the chosen value of fanout to ensure proper operation in a more complex system of gates.

5.4.1 Gate Transfer Function

This section calculates the transfer function of a single gate, from which small- and large-signal gain can be obtained. If the operation of all gates in the system were identical, then the large-signal gain (within the constraints of acceptable noise margin), would give the fanout of the gate. However, as shown in the following section on cascading, the transfer function of the first gate is different than the transfer functions of all subsequent gates. These subsequent transfer functions converge rapidly to the function that is used to define the final fanout and noise margins.

In this section, then, the large-signal gain serves as a basis for comparison of various gates based on the collision and dragging interactions. The gates with the best characteristics are then studied in cascaded operation in the following sections, for which the fanout is obtained more formally based on the (approximate) asymptotic transfer function of the cascaded system.

Since large-signal gain is of interest, the self-consistent operating points must shift (assuming that fanout equals large-signal gain), such that the input levels $I_L = O_L/G$ and $I_H = O_H/G$, where G is the desired large-signal gain. The noise margins are then defined

$$NM_L = P_{IL} - O_L/G = P_{IL} - I_L \quad (5.48a)$$

$$NM_H = O_H/G - P_{IH} = I_H - P_{IH}, \quad (5.48b)$$

where the definitions are now precise, as indicated in Figure 5.26. Notice that there is a tradeoff between large noise margin and large fanout, as will also be evident in the following discussions. Large-signal gain can be increased for high-contrast electronic logic gates through the use of level shifters, for which there is no direct analog in optics. This is the reason why the noise margins are defined as above for optical logic.

Collision Gate

The first collision gate studied is for $r = 1$, gate length $10 Z_0$, and in the absence of absorption. For the normalized angle $\kappa = 0.37$ and initial separation $r_s = 5$, the threshold contrast $\tau = 2.13$, suggesting that the use of this gate for restoring logic may not be possible. The optimal angle $\kappa = 0.11$ indicated in Figure 5.14 gives larger τ , but results in a non-monotonically decreasing transfer function and will not be considered. The transfer function for $\kappa = 0.37$ is shown in Figure 5.27, which illustrates that there are no valid logic states of the device, because small-signal gain is everywhere less than unity. Right at the threshold level of 0.77, the small-signal gain is 0.96. As a result, large-signal gain cannot be provided and this gate will not function as a logic device. This gate will function as a simple switching device without gain, however, with on-off ratio, or switching contrast, greater than 20.

This result for the $r = 1$ collision gate is not entirely unexpected. Figure 5.13 clearly shows that the threshold contrast criterion is not met for $r = 1$, because the pump undergoes very little angular deflection. Figure 5.13 also shows that the collision gate has very large threshold contrast when $r > 1$. Therefore, it is expected that these gates will function with small-signal gain. The specific case of $r = 3$ was examined in more detail, in Figure 5.18 from which the value $\kappa = 0.37$ is obtained, and in Figure 5.24 in the presence of absorption. Even including the effects of absorption, the threshold contrast for the $r = 3$ collision gate exceeds 10 for gate lengths of about $6 Z_0$ or longer, with $r(d) \approx 2$. It is therefore expected that in this range, restoring logic gates with large-signal gains of at least 2 can be obtained.

Figure 5.28 shows the transfer functions for the $r = 3$ collision gate of lengths $5 Z_0$ and $10 Z_0$. The threshold level of the $d = 5$ gate is 0.98, which is expected since the threshold contrast $\tau = 1.1$. The rule of thumb used in obtaining the minimum level of $\tau = 10$ was that $P_{OL} \approx 0.1$, and $P_{OL} = 0.11$ here, but given that $1/\tau = 0.91 \gg P_{OL}$, it is clear that the high input level $P_{IH} > 1$. But, it was mentioned earlier that large-signal gain is still obtained when $P_{IH} (= 1.23) < r(5) (= 2.14)$, as it is here, with value $G = 1.6$. Given that $P_{IL} = 0.63$, equation 5.22 suggests that $\tau \geq 0.56$ for at least unity small-signal gain. Since $\tau = 1.1$, an approximate lower bound is given by $1.1/0.56 = 1.9$. In fact, the small-signal gain of this gate is 4.71. Therefore, this

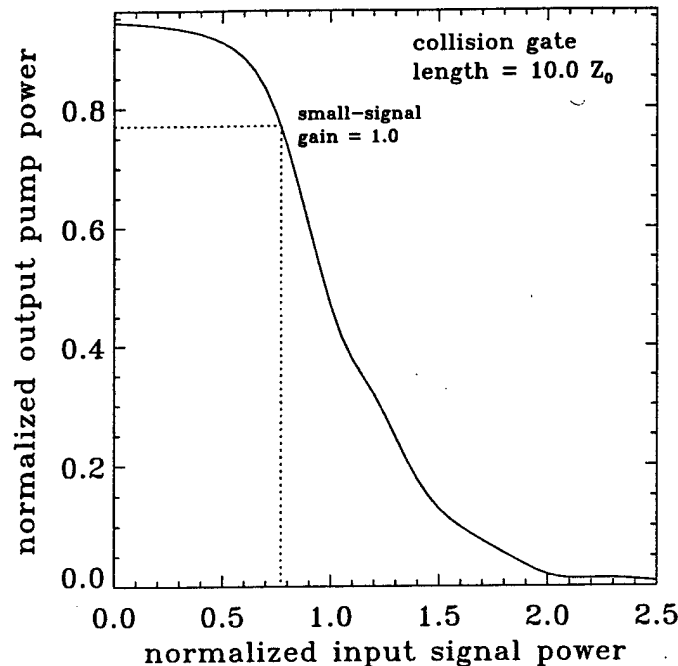


Figure 5.27: Transfer function for the $r = 1$ collision gate of length $10 Z_0$ with normalized angle $\kappa = 0.37$ and separation $r_s = 5$. Absorption is neglected and the threshold contrast $\tau = 2.13$. This gate does not provide large-signal gain.

gate provides one example where small- and large-signal gain are obtained when $\tau < 10$. Since the output saturates rapidly for input levels less than P_{IL} and greater than P_{IH} , this gate exhibits large noise margins.

The threshold contrast $\tau = 3.1 \times 10^4$ for the $d = 10$ gate indicates that the point $(1, 1/\tau)$ lies well within the saturated region of valid low states. Therefore, small-signal gain is guaranteed (because the normalized high output $r(10) = 2.11 > 1$) given a smooth transfer function (small-signal gain is 5.58 for this gate), and large-signal gain of at least $r(10) = 2.11$ is expected. For $G = 2.6$, the stable operating points lie well within the region of valid high and low device states (i.e. the regions where small-signal gain is less than unity), leading to noise margins of $NM_L = 0.18$ and $NM_H = 0.11$, which should be sufficient for sustained, cascaded, operation. Note that this only provides an indication of the cascability of these devices, because even though the output power of the device (which is divided by the fanout and serves as the signal input to the next gate) may correspond to a particular input power, the shape of the new input beam is not exactly the same as the original input signal soliton shape due to clipping at the output aperture. This effect is minimized in the absence of the signal such that the pump is undeviated, where symmetric clipping of the wings resulting in a small power loss occurs, but the overall effect is to alter the shape of the transfer characteristic. As a result, true logic-level restoration is not obtained for the first gate (because the shape is not restored), but it is obtained for cascaded operation, in which the clipped output is the same input for every gate, as examined in detail in section 5.4.2.

Dragging Gate

The comparisons of previous sections showed that the dragging gate generally has shorter length than the collision gate for a given value of r and τ , both with and without absorption, which suggests that the dragging gates can achieve greater large-signal gain at a fixed gate length. These properties are verified here.

Figure 5.29 shows the transfer functions for two otherwise identical $r = 1$ dragging gates of length $10 Z_0$ except that the bottom function is plotted including the effects of absorption with $s_{opt} = 0.2627$ at $10 Z_0$. In both cases, $\tau \approx 100$, which indicates that the threshold level < 1 , and $P_{OL} \approx 0.1$, which satisfies the rule of thumb employed to place the condition $\tau > 10$. Unlike the $r = 1$ collision gate, each $r = 1$ dragging gate provides large-signal gain, $G = 2.0$ in the case without absorption, and $G = 1.4$ in the case with absorption.

Even though large-signal gain (which should lead to fanout as well) can be obtained for the $r = 1$ dragging gates, it is desirable to obtain greater values for increased fanout. For general purpose logic, a minimum fanout of 2 is necessary. With the inclusion of material absorption, this may not be possible for the $r = 1$ dragging gate, for which the maximum large-signal gain

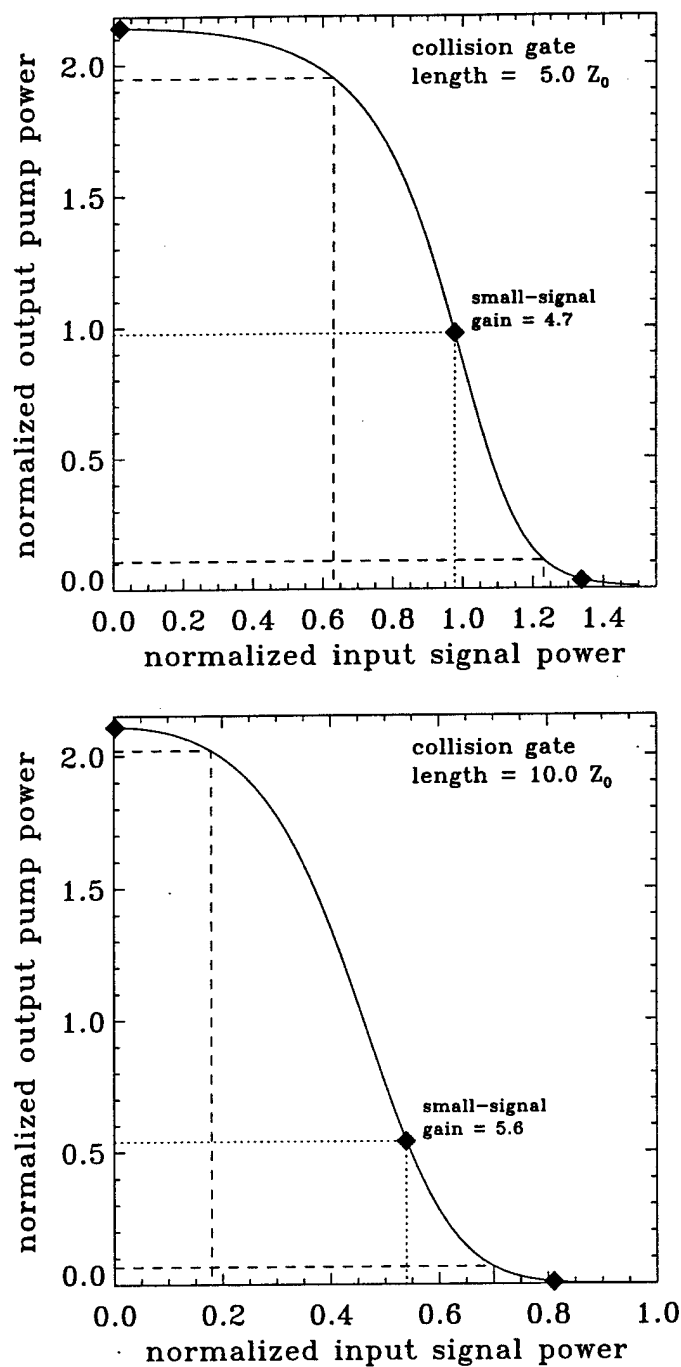


Figure 5.28: Transfer functions for the $r = 3$ collision gate of lengths $5 Z_0$ (top) and $10 Z_0$ (bottom) with normalized angle $\kappa = 0.37$ and separation $r_s = 5$. Absorption is included using $s_{opt} = 0.2627$ for each gate length. The threshold contrasts are $\tau = 1.10$ for the $5 Z_0$ gate length and $\tau = 3.1 \times 10^4$ for the $10 Z_0$ gate length. These gates provide large signal gains of $G = 1.6$ with $NM_L = 0.61$ and $NM_H = 0.11$ (top) and $G = 2.6$ with $NM_L = 0.18$ and $NM_H = 0.11$ (bottom). The high noise margins represent 8.2% and 14% deviation about I_H , respectively, and the filled diamonds denote the operating points and threshold.

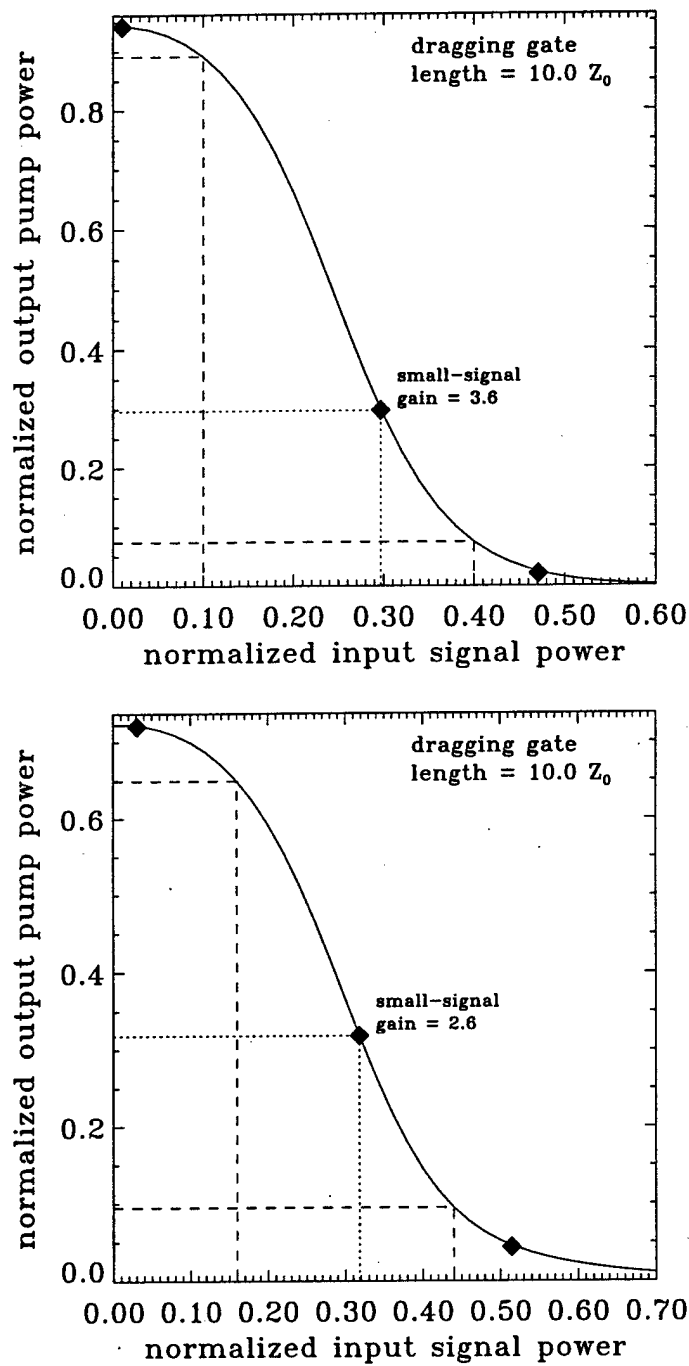


Figure 5.29: Transfer functions for the $r = 1$ dragging gate of length $10 Z_0$ with normalized angle $\kappa = 0.8$ and separation $r_s = 0$. Absorption is neglected in the top gate (with threshold contrast 150), but included in the bottom gate (with threshold contrast 131) using $s_{opt} = 0.2627$. These gates provide large signal gains of $G = 2.0$ with $NM_L = 0.09$ and $NM_H = 0.07$ (top) and $G = 1.4$ with $NM_L = 0.13$ and $NM_H = 0.07$ (bottom). The high noise margins represent 15% and 14% deviation about I_H , respectively, and the filled diamonds denote the operating points and threshold.

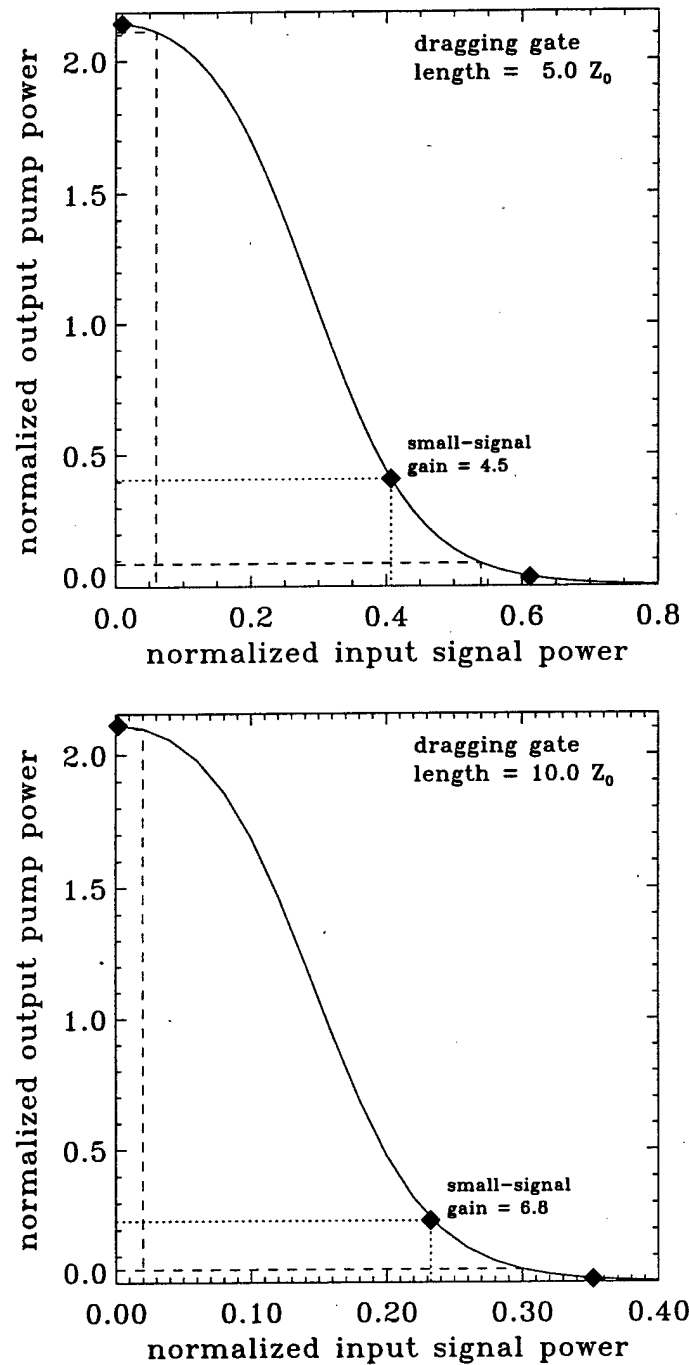


Figure 5.30: Transfer functions for the $r = 3$ dragging gate of lengths $5 Z_0$ (top) and $10 Z_0$ (bottom) with normalized angle $\kappa = 0.8$ and separation $r_s = 0$. Absorption is included using $s_{opt} = 0.2627$ for each gate length. The threshold contrasts are $\tau = 1.8 \times 10^3$ for the $5 Z_0$ gate length and $\tau = 7.4 \times 10^3$ for the $10 Z_0$ gate length. The gates provide large signal gains of $G = 3.5$ with $NM_L = 0.05$ and $NM_H = 0.07$ (top) and $G = 6.0$ with $NM_L = 0.02$ and $NM_H = 0.05$ (bottom). The high noise margins represent 12% and 15% deviation about I_H , respectively, and the filled diamonds denote the operating points and threshold.

is 1.4. Therefore, the $r = 3$ dragging gates (including absorption) are investigated in order to provide increased fanout. These gates are then examined in section 5.4.2 in cascaded operation in order to determine the actual fanout.

The transfer functions for the $r = 3$ dragging gates are shown in Figure 5.30, for gate lengths of $5 Z_0$ and $10 Z_0$, each with $s_{opt} = 0.2627$. Using the same device parameters, the dragging gates provide much greater large-signal gain, $G = 3.5$ for the $d = 5$ gate and $G = 6.0$ for the $d = 10$ gate, than the corresponding collision gates. Because the dragging interaction works so well, the allowable noise margin about the low input is much smaller than for the collision gates, which possess much broader

saturated regions at the high output level. Here, $P_{IL} = 0.06$ for $d = 5$ and $P_{IL} = 0.02$ for $d = 10$, resulting in low noise margins of $NM_L = 0.05$ and $NM_L = 0.02$ for $I_L = 0.01$ and $I_L = 0.0$, respectively, and shows that the dragging gates intrinsically provide higher gain with smaller NM_L (because P_{IL} are so small), and comparable high noise margin NM_H (on a percentage basis).

The following section of this chapter consider only the $r = 3$ collision and dragging gates in the presence of absorption with $s_{opt} = 0.2627$ for gate lengths of $5 Z_0$ for dragging and $10 Z_0$ for collision and dragging. It is useful at this point to consider the interaction parameters, such as gate length, beam widths, and beam powers, in real units. Since $s_{opt} = 0.2627$, and it was assumed that $\alpha_0 = 0.1 \text{ cm}^{-1}$ for fused silica waveguides, the gate length in real units is given by $\alpha_0 z = 0.2627$, or $z = 2.63 \text{ cm}$. For the $d = 5$ gate length, $w_0 = 23.9 \text{ } \mu\text{m}$ (FWHM $42.2 \text{ } \mu\text{m}$), and $w_p = w_0/r = 7.97 \text{ } \mu\text{m}$ (FWHM $14.0 \text{ } \mu\text{m}$). With $n_2^I = 3.3 \times 10^{-16} \text{ cm}^2/\text{W}$, the peak intensities of the fundamental solitons are $I_0 = 2.23 \times 10^{11} \text{ W/cm}^2$ and $I_p = 2.01 \times 10^{12} \text{ W/cm}^2$. Assuming a $4 \text{ } \mu\text{m}$ thick waveguide, the corresponding fundamental soliton powers are $P_0 = 4.27 \times 10^5 \text{ W}$ and $P_p = 12.8 \times 10^5 \text{ W}$. Note that the actual switching power is less than P_0 . For the $d = 5$ dragging gate, the switching power is $I_H P_0 = 0.61 P_0 = 2.6 \times 10^5 \text{ W}$.

The length of the $d = 10$ gates is $z = 2.63 \text{ cm}$ as well, with $w_0 = 16.9 \text{ } \mu\text{m}$ (FWHM $29.8 \text{ } \mu\text{m}$) and $w_p = 5.63 \text{ } \mu\text{m}$ (FWHM $9.93 \text{ } \mu\text{m}$). The corresponding peak soliton intensities are a factor of 2 greater than for the $d = 5$ gate, while the powers greater by a factor $\sqrt{2}$, or $P_0 = 6.05 \times 10^5 \text{ W}$, while $P_p = 18.1 \times 10^5 \text{ W}$. The switching power for the collision gate is $0.81 P_0 = 4.9 \times 10^5 \text{ W}$ and $0.35 P_0 = 2.1 \times 10^5 \text{ W}$ for the dragging gate. It is interesting to note that the switching power of the dragging gate is less than half that of the collision gate, and less than even that of the $d = 5$ dragging gate.

5.4.2 Ring Oscillator Configuration

The previous section examined the transfer functions for single-stage logic gates and showed that sufficient large-signal gain exists for the $r = 3$ collision gate of length $d = 10$ and for the $r = 3$ dragging gates of lengths $d = 5$ and $d = 10$ that fanouts of 2 or greater should be obtainable. It is the purpose of this section to obtain the fanout values for these gates in the ring oscillator configuration. As mentioned previously, the maximum fanout is typically determined by the minimum allowable noise margins. Thus, for a given value of fanout, the gate transfer function is recalculated and new noise margins obtained.

In cascaded operation, the output of one gate serves as the new signal input to the next gate. For an infinite cascade of inverters (i.e. a single gate with feedback), the result is a ring oscillator. Since an infinite simulation is not possible, the cascade of three gates is used, and it is shown that the transfer function of the third gate is almost identical to the second gate, indicating that rapid convergence is obtained and that a deeper cascade need not be considered. These transfer functions are consistent in that the large-signal gain based on the stable operating points is equal to the fanout. The fanout is not necessarily equal to the maximum obtainable large-signal gain calculated from all possible operating points within the valid high and low operating regions, because the maximum value may result in unacceptable noise margins.

Unlike for electronics, the transfer functions of the second and third gates are generally different than that of the first gate because the first gate does not provide complete logic level restoration of all of the multi-dimensional beam parameters due to clipping by the output aperture. However, because of the presence of a fresh pump at each gate, the output of the second gate (divided by the fanout factor) is nearly identical to the signal input to that same gate. For gates farther down the line, the output converges to the exact form of the input, and complete logic level restoration is obtained. This convergence rate of this process is fast enough (as indicated by the convergence of the transfer functions) that for the present purposes, the second gate is said to have complete logic level restoration.

The pump output of one gate can be used as a signal input to subsequent gates by simply flipping the polarization and tilting the angle. Changing the state of polarization is not even necessary, because the definition of pump and signal polarizations can be exchanged at every gate in the cascade. Figure 5.31 shows one simple implementation of the cascading of two gates, and is easily extendible to include more gates. A telecentric imaging system (based on lenslet arrays for spatially parallel operation) performs 1:1 imaging between the output of the first gate and the input to the second gate. Note that magnification can be used to scale the width of the new signal to that of the original signal in the first gate, but, due to broadening of the pump in the presence of absorption, this is not necessary and will not be employed. A tilted partial beam splitter placed just before the polarizing beam splitter serves two purposes. The first purpose is to split off the pump output to fanout to other gates. The second purpose is to laterally offset the pump so that off-center propagation through the second lens will result in an angle change equal to that of the signal in the first stage. The polarizing beam splitter allows the lossless overlap of the pump and signal and also serves to block any signal leakage from the previous stage.

The cascaded transfer functions for the $r = 3$ collision gate are shown in Figure 5.32 for fanout of 3. Because the maximum input to a cascaded device is given by the maximum output O_{max} of the previous device divided by the fanout, the maximum cascaded input level is given by $I_{max} = O_{max}/F$, where F is the fanout factor. Therefore, the maximum high noise margin is $NM_H = I_{max} - P_{IH}(F)$, where the dependence of P_{IH} on the fanout factor is included because the transfer curve changes with F . Likewise, the maximum low noise margin is $NM_L = P_{IL}(F)$, because the minimum possible low output level is $O_{min} = 0$.

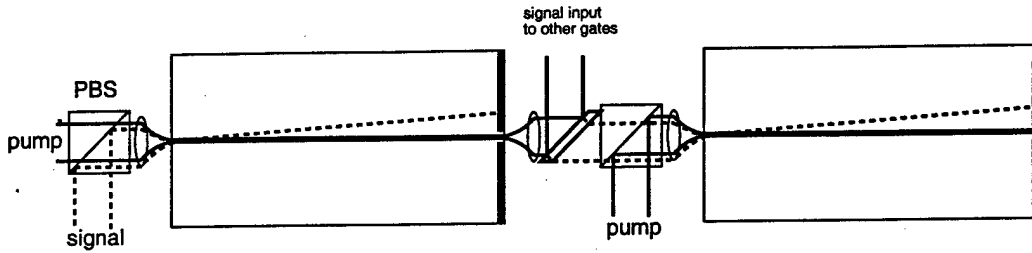


Figure 5.31: Cascaded spatial soliton angular deflection gates. A free-space region between the nonlinear sections allows the fresh pump and cascaded signal to be brought into coincidence without nonlinear interaction. With high space-bandwidth product free-space interconnects, the large degrees of spatial parallelism inherent in this geometry can be utilized to implement complex switching fabrics.

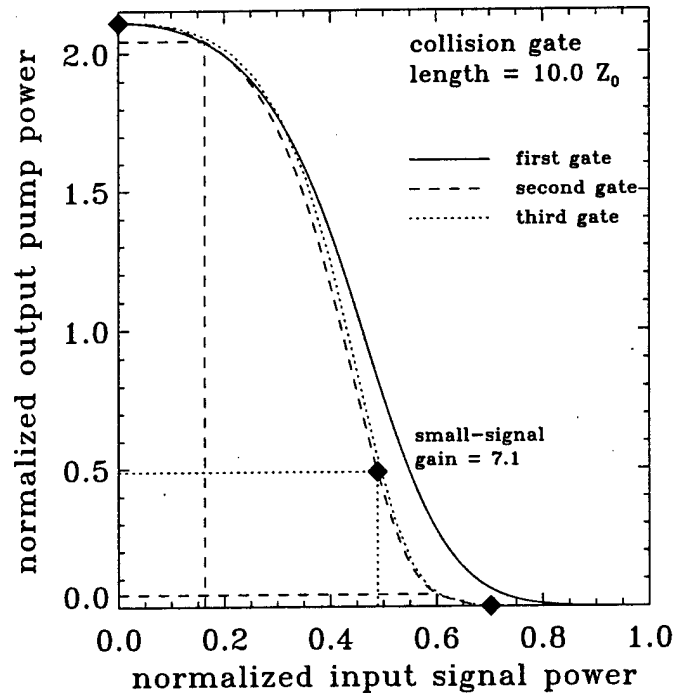


Figure 5.32: Transfer functions for the cascaded $r = 3$ collision gate of length $10 Z_0$ with normalized angle $\kappa = 0.37$ and separation $r_s = 5$. Absorption is included with $s_{opt} = 0.2627$. With a fanout of 3, the operating points lie well within the saturated levels and noise margins of $NM_L = 0.16$ and $NM_H = 0.10$ are obtained, where the high noise margin represents 14% deviation about $I_H = 0.70$. The filled diamonds denote the operating points and threshold.

In order that the large-signal gain equals the fanout ($G = F$) for self-consistent operation, $I_H = O_H/G$ must be the high input level corresponding to the operating point of the low output state, and $I_L = O_L/G$ must be the low input level corresponding to the operating point of the high output state. In the example of Figure 5.32, $O_H \approx 0$, such that the noise margins of this gate are then the maximum noise margins just discussed.

The transfer functions for the second and third gates are nearly identical, indicating that complete logic-level restoration has been obtained at the second gate. Therefore, stable cascaded operation with fanout of 3 is demonstrated, which will function indefinitely (within the allowable noise margins). The transfer function sharpens up for the second and third gates, which have greater small-signal gain (7.1) than the first gate (5.6). The reason for this has to do with the clipping produced by the output aperture, and is also the reason why greater fanout can be achieved than the large-signal gain of the first device would suggest. For high output states of the first device, very little clipping of the pump occurs, which serves as a "whole" signal input (divided by the fanout factor) to the second gate, resulting in strong deflection of the next clean pump. As the first gate starts to switch into the low state for inputs greater than P_{IL} , the pump becomes clipped, and results in less deflection of the pump in the next

stage than the power of the resultant cascaded signal would suggest (because this new signal does not have the same shape as the original signal in the first gate for which the input power scale is defined). This effect creates a sharper transition region between the high and low states for the second and all subsequent devices, which also reduces the input switching level P_{IH} , resulting in larger gain. Since the clipped input to the third gate looks like the input to the second gate, complete logic level restoration is obtained and the transfer functions converge.

Figure 5.33 shows the same information for the $r = 3$ dragging gate of length $5 Z_0$. The top plot is for fanout of 3.0, and the bottom plot is for fanout of 5.0. Comparison of these plots illustrates the expected tradeoff between fanout and noise margin. The transfer function for the second and third gates in the fanout 3 case shows that the high input operating point $I_H = 0.71$, which is the maximum output of the gate of 2.14 divided by the fanout factor. The only way to access input levels greater than this value is through noise processes. Therefore, the single-sided noise margin is given by $0.71 - P_{IH} = 0.07$ in normalized power units, where P_{IH} is the minimum value of the high input level that lies within the region of small-signal gain less than unity.

For the fanout 5 gate, the high input operating level is $I_H = 0.43$, which lies outside the region of small-signal gain < 1 . Therefore, this gate is said to have negative noise margin and cannot operate in a stable, cascaded system due to the lack of a directly accessible saturated input level region about the low output state O_L . Note that the transfer functions for the second and third gates are nearly identical, indicating that logic-level restoration is provided in this situation as well. These two fanout examples illustrate the tradeoff between fanout and noise margin. The critical fanout at which zero noise margin is obtained is about 3.9; below this fanout, the noise margin increases positively, above this fanout, the noise margin becomes more negative.

Finally, the high-gain $r = 3$ dragging gate of length $10 Z_0$ is examined in Figure 5.34 for fanout of 5. This choice of fanout results in a self-consistent system with noise margin about the high input level of 0.05, which allows for about 12% variation in the negative direction (i.e. loss) due to noise. Note that the high-input noise margin in the positive direction (i.e. for gain about I_H) is for all practical purposes unbounded (as it is for all gates in this section) which allows for highly asymmetric distribution of signal levels; therefore, only the more restrictive single-sided noise margin is presented.

An interesting comparison can be made between the optical logic gates studied here and electronic logic gates. For an electronic gate of fixed geometry, the switching time is proportional to the fanout, because the single gate must charge the input capacitance of all subsequent gates. The optical logic gates in this section are all of the same length, and hence all have the same latency or switching time, and the fanout factor is determined by the values of normalized gate length d and initial power ratio r . As shown here, fanout of 5 can be obtained with a $d = 10$ dragging gate with the same latency as the $d = 5$ dragging gate which produces a fanout of 3. This behavior can be extended to obtain greater fanouts using larger values of d and r within the limits placed by absorption. The unique scaling of these optical gates could be capitalized on at the systems level through the use of very high fanout devices.

This section has shown that the $r = 3$ collision and dragging logic gates can be cascaded in an indefinite manner (i.e. asymptotically stable) with fanouts of 3 or greater (where the minimum necessary for the implementation of arbitrary logic is 2) in the presence of realistic absorption parameters. In addition, these gates possess true logic-level restoration which converges rapidly in the first few cascaded stages.

5.4.3 Two-Input NOR Configuration

The final cascaded configuration studied is the two-input NOR, or 2-NOR, gate. A NOR gate is one example of a single member logically-complete family. Thus, its implementation is of utmost importance in order to demonstrate the ability to implement arbitrary logic functionality. In a cascaded 2-NOR gate, two signals interact with a single pump in two successive stages, so that either one can switch the output to low. Therefore, the pump output of the first stage is used as the pump input to the second stage. The pump output (divided by the fanout factor) of the second stage, which is the output of the NOR gate, is then used as the signal input to a second gate (which can be a single inverter or another NOR gate for which there are two signal inputs) in order to determine the fanout of the NOR gate when driving a subsequent gate. There are three situations for which the transfer function of the 2-NOR gate needs to be calculated: only signal 1 present, only signal 2 present, and both signal 1 and signal 2 present. The latter case does not differ significantly from the first case and will not be considered.

The first geometry considered here is illustrated in Figure 5.35, and is the direct analogy of one possible fiber implementation with temporal soliton dragging [66]. Since stable, cascaded, operation is of interest here, the 2-NOR gate is studied using the output of an inverter (with fanout factor of 3) as either of the signal inputs. The output of the 2-NOR gate is then cascaded into the signal input of a subsequent inverter or into one of two inputs to another 2-NOR gate. This second cascading stage determines the allowable fanout of the 2-NOR gate. The inverter and each stage of the NOR gate are based on the $r = 3$ dragging interaction with gate length $5 Z_0$ and $s_{opt} = 0.2627$. The inverters are identical to those studied previously, while the output aperture sizes of the two stages of the 2-NOR gate need to be considered more closely. Since the pump passes through two absorbing stages, the beam width at the output of the 2-NOR gate (i.e. the output of the second stage) is wider than at the

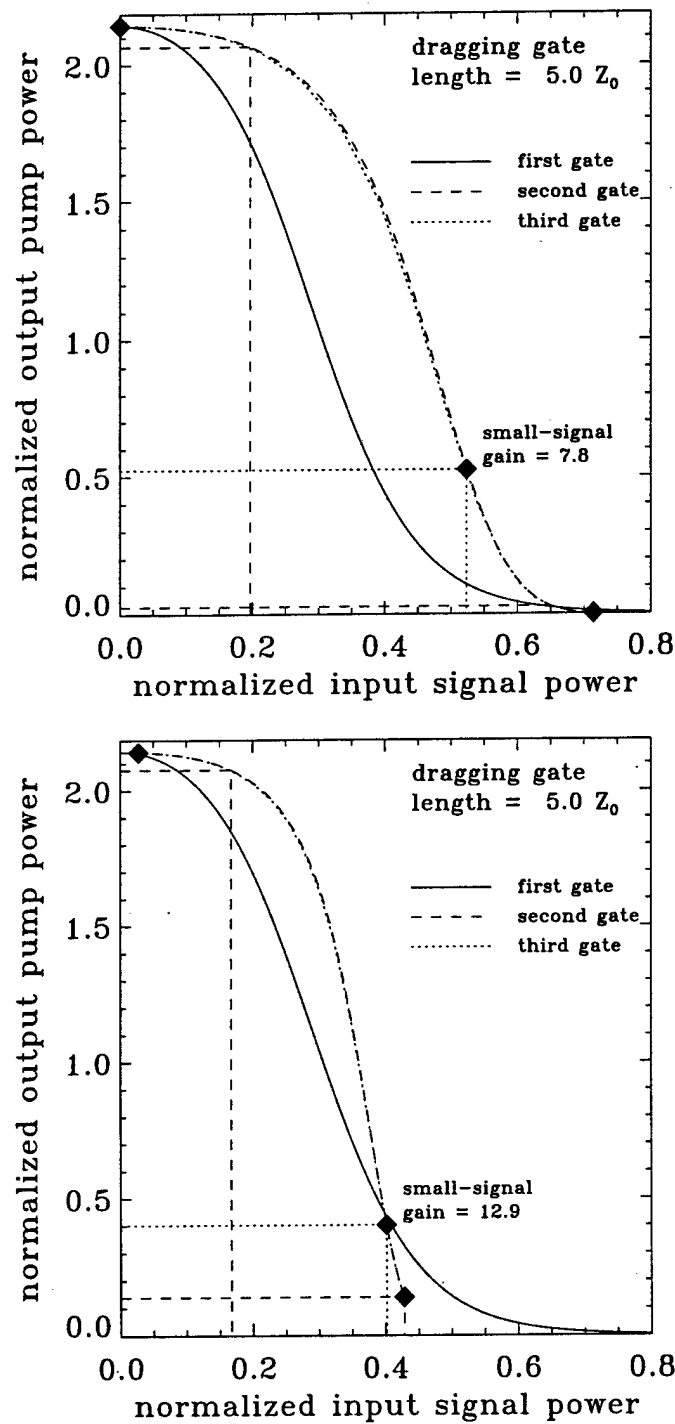


Figure 5.33: Transfer functions for the cascaded $r = 3$ dragging gate of length $5 Z_0$ with normalized angle $\kappa = 0.8$. Absorption is included with $s_{opt} = 0.2627$. The transfer functions are plotted for fanout of 3 (top) and 5 (bottom). The top gate provides large signal gain of $G = 3.0$ with $NM_L = 0.196$ and $NM_H = 0.07$, which represents 9% deviation about I_H . Since the operating point of the bottom gate lies in the region of small-signal gain, this gate has negative noise margin about I_H . The filled diamonds denote the operating points and threshold.

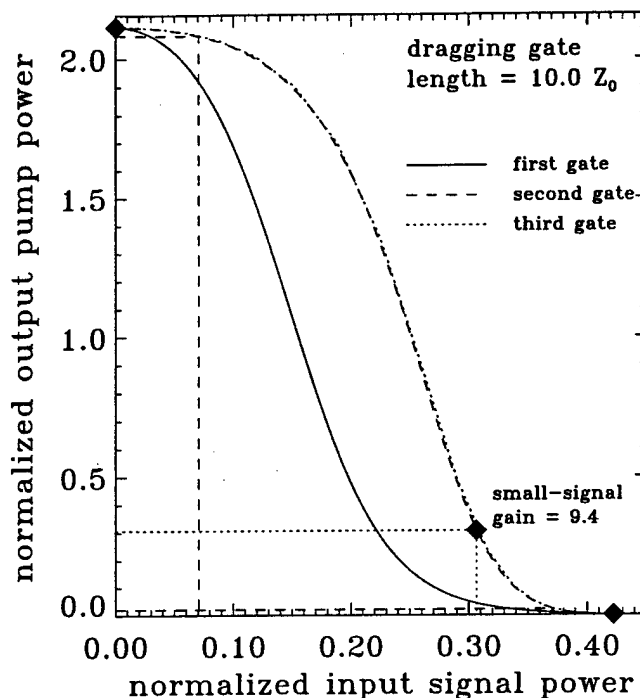


Figure 5.34: Transfer functions for the cascaded $r = 3$ dragging gate of length $10 Z_0$ with normalized angle $\kappa = 0.8$. Absorption is included with $s_{opt} = 0.2627$. With a fanout of 5, the operating points lie well within the saturated levels and noise margins of $NM_L = 0.07$ and $NM_H = 0.05$ are obtained, where the high noise margin represents 12% deviation about $I_H = 0.42$. The filled diamonds denote the operating points and threshold.

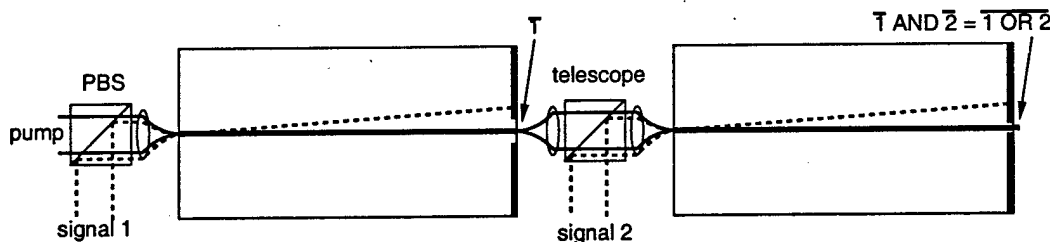


Figure 5.35: Cascaded spatial soliton dragging gates implementing a two-input NOR gate, in which the pump is shared between two symmetric inverter stages. A free-space region between the nonlinear sections allows the pump and signal 2 to be brought into coincidence without nonlinear interaction. With high space-bandwidth product free-space interconnects, the large degrees of spatial parallelism inherent in this geometry can be utilized to implement complex switching fabrics.

output of the single-stage inverter. In all cases, the gate output aperture is set to 3.5 times the width of the pump at the output $w_p(d)$, where $d = 5$ for the inverter and $d = 10$ for the symmetric 2-NOR, such that 94.2% of the power of the undeviated pump passes.

Now the question arises whether to place an aperture at the output of the first stage of the 2-NOR gate. If there were no aperture on the first stage, then, neglecting the effects of the imaging system for the moment, the permanent angular deviation of the pump would integrate over the second stage into additional spatial shift. Thus, the aperture of the second stage serves as the discriminating element for both signal 1 and signal 2 inputs. The imaging system significantly modifies this behavior, however, but results in the same net effect. Perfect telescopic imaging, as assumed in Figure 5.35 and shown in more detail in Figure 5.36, imparts no phase aberration, and, if unity-conjugate imaging is used, no magnification. The spatial coordinates in the image plane are inverted with respect to the coordinates in the object plane, such that spatial deviation r of the pump at the output of the first stage is transferred to the input of the second stage as the spatial deviation $-r$. In addition, during the imaging operation, the propagation angle of the pump is reversed as well, so that a deflection angle of θ is turned into an angle

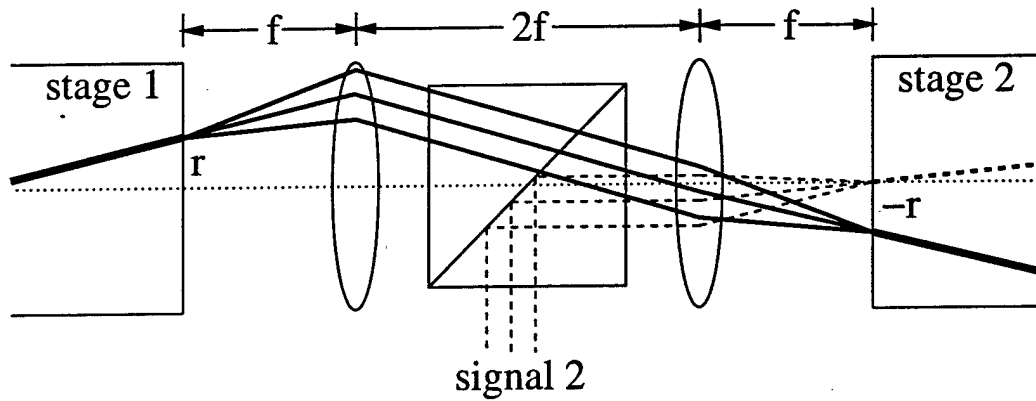


Figure 5.36: Telescopic imaging system between the two stages of the 2-NOR gate. A pump that is undeviated in the first stage images symmetrically about the optical axis to the second stage (not shown). A pump that is deviated in the first stage (solid lines) in both spatial position and angle, is imaged to the conjugate plane shifted in the opposite direction and with opposite angle. The signal 2 input is indicated by the dashed lines, and would spatially overlap the undeviated pump at the input to the second stage.

$-\theta$ at the input to the second stage. These effects mean that the spatial shift accumulated in the first stage is transferred to the second stage and the angular deviation of the pump incurred in the first stage leads to additional spatial shift in the second stage. Hence, angular deflection of the pump in the first stage is leveraged over the additional length of the second stage, effectively doubling the device length.

In the simulations presented here, the output aperture size for the first stage is set at $17.5 w_p(5)$, or 5 times the width that would be used if this stage were a stand-alone inverter gate. The main reason for widening the aperture is to minimize the diffractive effects on the pump, so that an undeviated pump will pass through the aperture with no clipping or power loss. This is important because the effects of the imaging optics between stages, which would low-pass filter and eliminate the sharp edges produced by the aperture, is not simulated. In fact, the aperture could be removed completely with the advantages of additional spatial shift in the second stage due to deflection in the first stage as discussed in the previous paragraph, but is used to block a strongly deviated pump (which would not overlap the signal in the next stage anyway), in order to keep stray pump light from bouncing off the boundaries of the computational grid in the second stage simulation.

The first case considered is the cascading of the symmetric 2-NOR gate into an inverter. The resulting transfer functions are shown in Figure 5.37. The most important difference between the transfer function of the symmetric 2-NOR gate (dashed curve) and the transfer function of the inverter is that the maximum output level of the 2-NOR is reduced. Even though both gates use the standardized $r = 3$ pump, the pump of the symmetric 2-NOR gate is used in two subsequent stages (each of the same length as the inverter) and incurs nearly twice the absorption. As a result, the maximum fanout of the 2-NOR is less than that of the inverter. Here, $F = 2.25$ is used so that the maximum gate output level of 1.63 divided by F gives similar signal input levels to a subsequent gate as would be produced by the inverter with fanout 3, thus ensuring consistent operation throughout. With this input level, the worst-case noise margins (resulting from the signal 2 input) for the 2-NOR are $NM_L = 0.15$ and $NM_H = 0.04$, where the latter represents 5.4% single-sided deviation about $I_H = 0.72$. The high noise margin can be increased by using a smaller fanout factor for the driving gate, but the resulting high input signal levels would be inconsistent with those produced by the $F = 3$ inverter. The use of different high input levels for the inverter and NOR gates does not necessarily result in any disadvantages, and in fact occurs with the use of the asymmetric 2-NOR geometry discussed later in this section, and is the approach taken in section 6.3.2 for the cascading between the spatio-temporal inverter and 2-NOR gates.

The two transfer functions (for each of the two signal inputs) of the 2-NOR gate differ considerably. The top plot of Figure 5.37 shows the transfer function when signal input 1 is used. In this situation, the effective gate length is $10 Z_0$ because any deflection produced in the first stage results in additional spatial shift in the second stage because of the widened aperture. Therefore, the minimum valid high input level $P_{IH} = 0.51$ for the first stage is much smaller than that of the second stage, where $P_{IH} = 0.68$. Alternatively, if the aperture width of the first stage were set to $3.5 w_p(5)$, then the clipped pump would propagate linearly and diffract in the second stage, again resulting in lower P_{IH} . In the absence of signal 1, the second stage works much like the single-input inverter, except for the fact the the pump propagates $5 Z_0$ with absorption before reaching the second stage of the NOR gate, and is therefore wider and carries less power. Due to additional absorption in the second stage, the pump is wider at the output of the 2-NOR gate than at the output of the inverter, so that greater spatial shift is necessary to produce a low output state. The result is that $P_{IH} = 0.68$ using the signal 2 input to the NOR gate, versus $P_{IH} = 0.65$ for the inverter. The

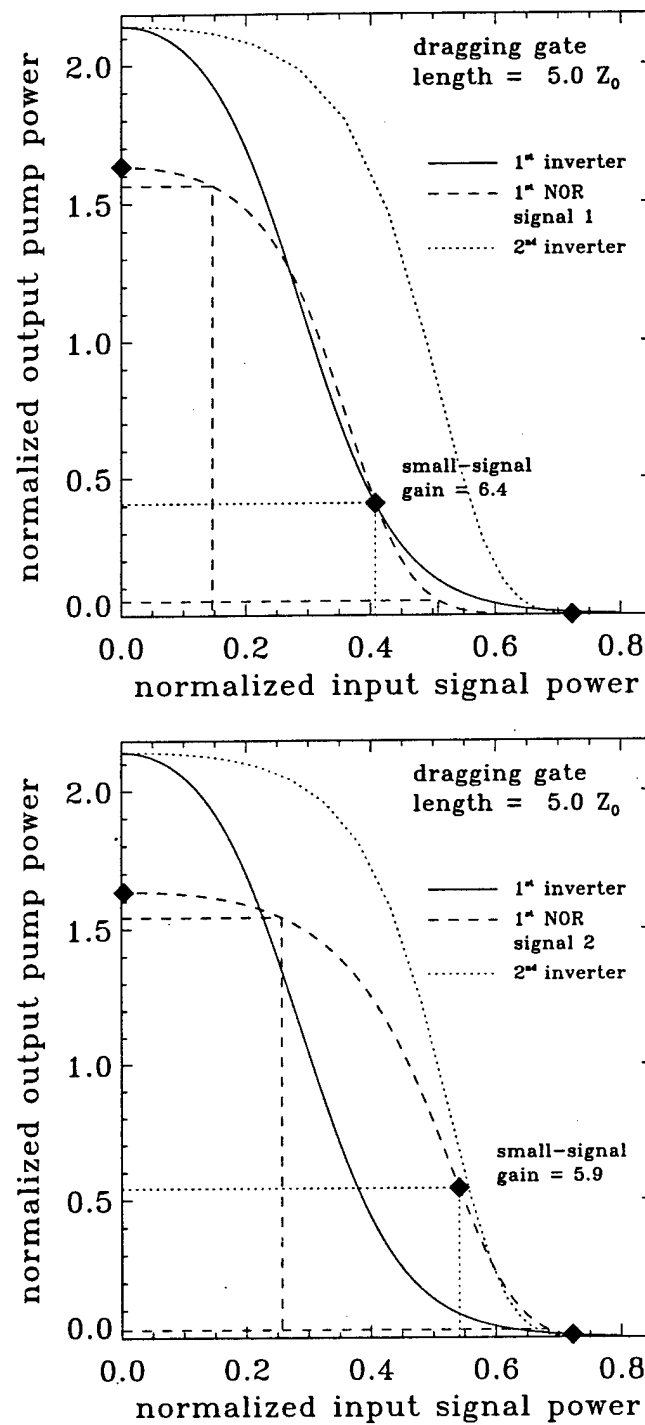


Figure 5.37: Transfer functions for cascaded symmetric 2-NOR and inverter gates based on the $r = 3$ dragging gate of length $5 Z_0$ with normalized angle $\kappa = 0.8$, and $s_{opt} = 0.2627$. The signal 1 (top) or signal 2 (bottom) input of the 2-NOR gate is driven by an inverter with fanout 3, and the signal input to the second inverter gate is driven by the output of the 2-NOR with fanout 2.25. The first case results in noise margins $NM_L = 0.15$ and $NM_H = 0.21$ (30% deviation about I_H), and the second case results in $NM_L = 0.25$ and $NM_H = 0.04$ (5.4% deviation).

worst-case transfer function for the 2-NOR gate is that produced by the signal in the second stage acting alone. It is this transfer function that is used when determining the noise margins and fanout.

Figure 5.37 also shows the transfer function of the inverter (dotted curves) when driven by the 2-NOR gate with $F = 2.25$. The two transfer functions of the second inverter are nearly identical, indicating that the cascaded operation between the 2-NOR gate and the inverter is relatively insensitive to the combinations of signal inputs to the NOR gate. These curves are different than that produced by the cascading of inverter gates alone, however, but the difference lies mainly in the saturated region about the high output state, which is broader in the current case, but has very little effect on the operation of the gate. For this inverter transfer function, $P_{IH} \sim 0.65$, leading to $NM_L \sim 0.06$, which is about 9% of $I_H = 0.71$. These numbers are essentially identical to those of the cascaded inverter shown in the top of Figure 5.33, and it is therefore expected that an inverter driven by a 2-NOR gate will cascade indefinitely as well.

Figures 5.38 and 5.39 show the transfer functions for a two-gate cascade of 2-NOR gates. The first two plots are for the signal 1 (first stage) input to the first 2-NOR gate and the second two plots are for the signal 2 (second stage) input. The most important aspects to notice from these series of plots is that the 2-NOR transfer functions converge in the situations when the corresponding inputs of the two gates are excited (i.e. the top plot of Figure 5.38 and the bottom plot of Figure 5.39) indicating complete logic-level restoration, and that all transfer functions lie within the boundaries of the worst-case provided by excitation of the second signal input.

As indicated in the previous series of plots for the cascading of the 2-NOR gate with the inverter, the 2-NOR gate provides fanout of 2.25 into an inverter gate, and the worst-case noise margins are $NM_L = 0.15$ (for signal input 1) and $NM_H = 0.04$ (for signal input 2), with $I_H = 0.72$ (which is produced by an inverter with $F = 3$ or by a NOR gate with $F = 2.25$). The figures here show that these performance parameters can be maintained in cascaded operation with another 2-NOR gate as well, such that a complete logic family with consistent operating parameters is obtained. The smallest low noise margin, $NM_L = 0.15$, is obtained from using the signal 1 input, because the extra gate length allows for reduction in the angular shift, so that smaller signal inputs can produce a resolvable spatial shift. The smallest high noise margin, $NM_H = 0.04$, is obtained by using the signal 2 input, as discussed before.

An alternative physical implementation of the 2-NOR gate results in standardized gate output levels, such that fanout is associated with the load (where the inverter and NOR gates have different I_H), rather than with the actual gate as in the previous situation (where the inverter and NOR gates have different high output levels, but consistent I_H). This implementation uses two separate stages as well, but keeps the total gate length at $5 Z_0$, the same length as the single-stage inverter. The main idea behind this configuration is that the bulk of the angular deflection of the pump is induced within the first one or two Z_0 of the material. As a result, the first stage can be very short, over which angular deflection occurs, with most of the spatial shift accumulated over the length of the longer second stage. Therefore, for the signal 1 input, the gate will operate similarly to the $d = 5$ inverter. The second stage must be long enough such that large accumulated spatial shift can be obtained for the signal 2 input. Similar to the previous symmetric geometry, the signal 2 input provides the worst-case gate performance here as well, and is the only situation studied for brevity.

The geometry of this asymmetric gate is shown in Figure 5.40. For the studies presented here, the length of the first stage is chosen to be $2 Z_0$, and the length of the second stage $3 Z_0$. Note from Figure 5.24 that the $r = 3$ dragging interaction produces large threshold contrast even down to about $2.5 Z_0$ gate length (indicating that angular deviation occurs over a very short distance), but the collision interaction generally requires much longer gate lengths. Because the solitons do not initially overlap in the collision interaction, the effective interaction distance over which angular deflection occurs is much longer than for dragging, such that the length of the first stage would be about $4 Z_0$ instead. Therefore, this asymmetric geometry and its extension to higher values of fan-in is used most effectively with the dragging interaction.

Two sets of cascaded transfer functions using the asymmetric 2-NOR gate are shown in Figure 5.41. The signal 2 input (which is again the worst case) of the NOR gate is driven by the output of the single stage inverter with fanout of 2.25. The fanout is reduced in order to increase the signal input level ($I_H = 0.95$) to compensate for the shorter length of the second 2-NOR stage. Because the maximum output level of this asymmetric NOR gate is the same as the inverter, it can drive the input of a cascaded inverter with fanout of 3.0 and either input of another 2-NOR with fanout of 2.25. The transfer functions for these two situations are shown in the top and bottom plots of the figure, respectively. As expected, the transfer function of the cascaded inverter is nearly identical to that of cascaded $r = 3$ dragging inverter at the top of Figure 5.33, and the transfer function of the second cascaded 2-NOR gate converges to that of the first 2-NOR. These results show that the output of the asymmetric 2-NOR is able to drive the inputs of an inverter or another 2-NOR in the same manner as the inverter, and that the fanout of either gate is determined by the load.

Larger fanouts than those shown here can be obtained by using the $r = 3$ dragging gate of length $10 Z_0$ as the basis for the 2-NOR gate. It should also be noted that even greater fanouts for the inverter and 2-NOR gate result by using larger values of r . These gates would need to be used in more complex logic situations where fanouts of 3 for the inverter and 2 for the 2-NOR are not sufficient. Another interesting consideration is the mixing of different gate families (such as an $r = 3$ gate with an $r = 4$

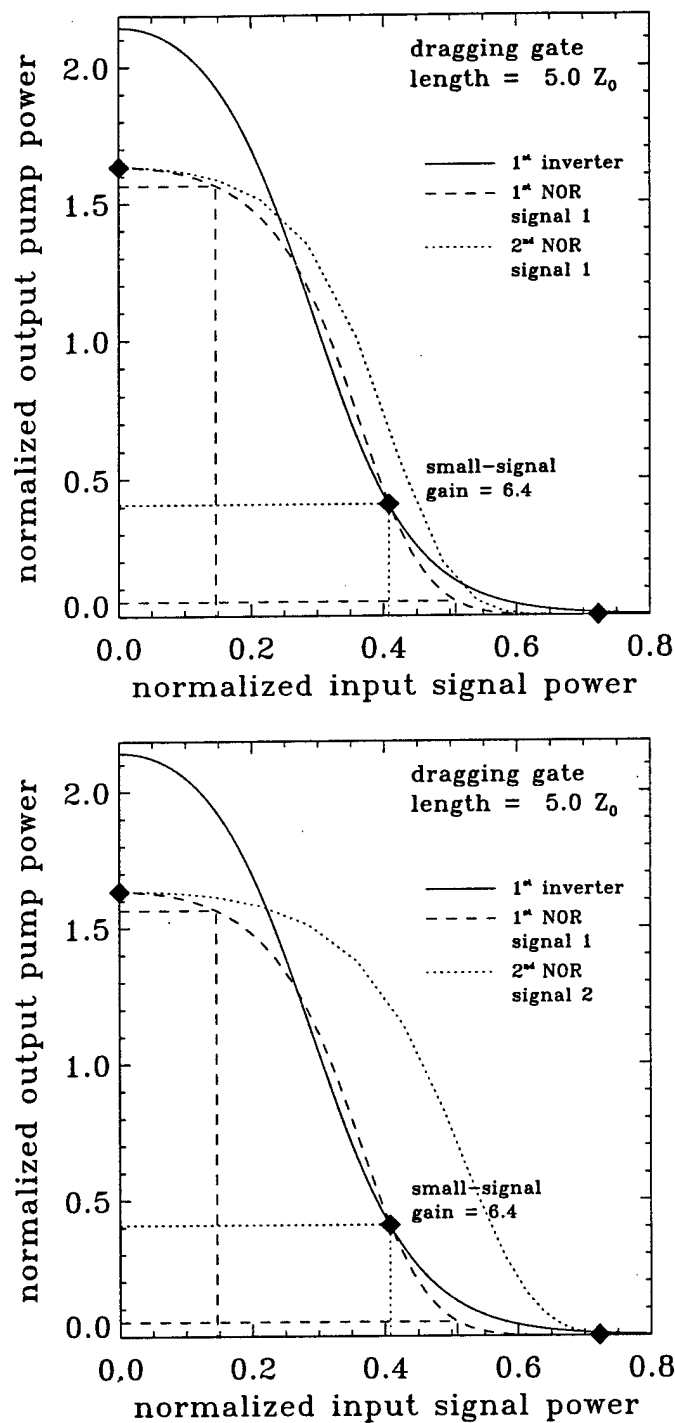


Figure 5.38: Transfer functions for cascaded symmetric 2-NOR gates based on the $r = 3$ dragging gate of length $5 Z_0$ with normalized angle $\kappa = 0.8$. The signal 1 input of the first 2-NOR gate in the cascade is driven by an inverter with fanout 3, and the first (top) or second (bottom) signal input to the second 2-NOR gate is driven by the first with fanout 2.25. The top situation results in noise margins $NM_L = 0.15$ and $NM_H = 21$, which represents 30% deviation about I_H , for the first NOR gate. The bottom situation is the same.

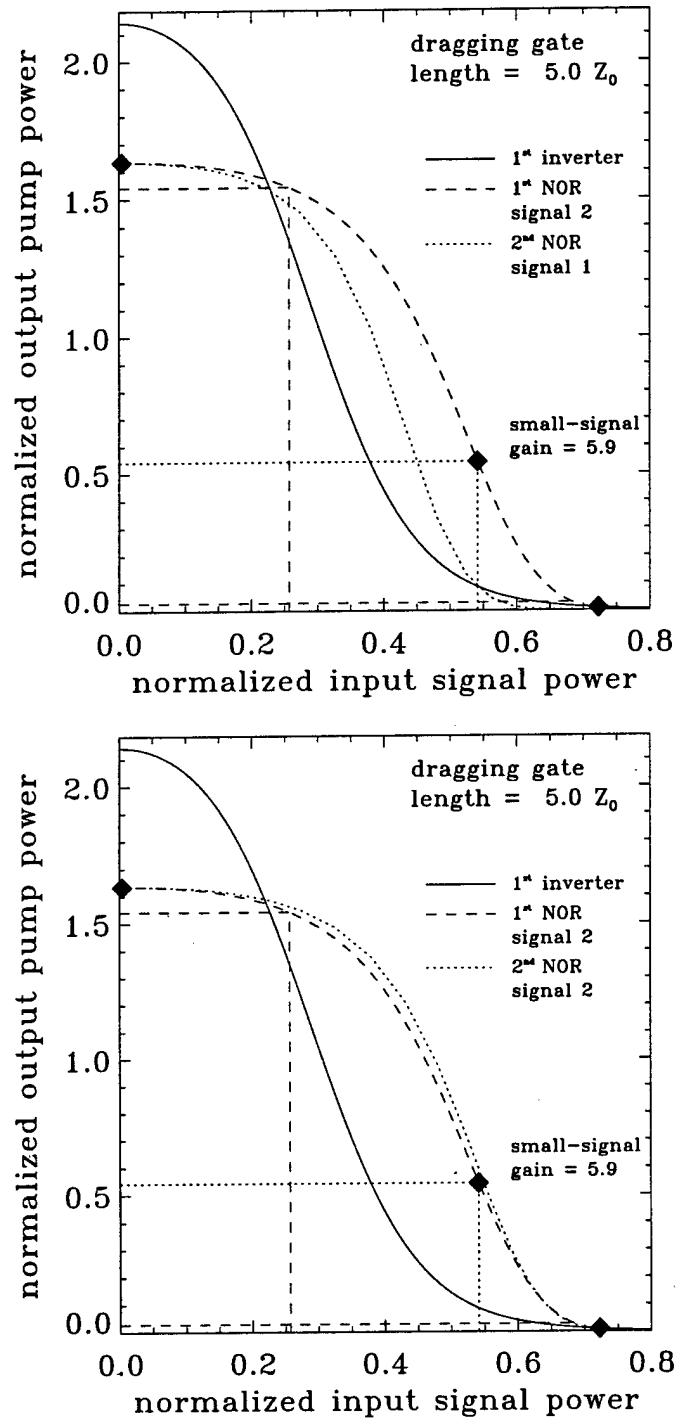


Figure 5.39: Transfer functions for cascaded symmetric 2-NOR gates based on the $r = 3$ dragging gate of length $5 Z_0$ with normalized angle $\kappa = 0.8$. The signal 2 input of the first 2-NOR gate in the cascade is driven by an inverter with fanout 3, and the first (top) or second (bottom) signal input to the second 2-NOR gate is driven by the first with fanout 2.25. The top situation results in noise margins $NM_L = 0.25$ and $NM_H = 0.04$, which represents 5.4% deviation about I_H , for the first NOR gate. The bottom situation is the same.

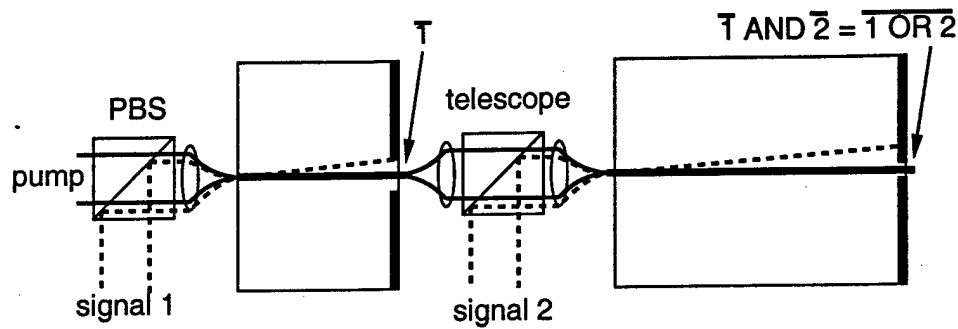


Figure 5.40: A two-input spatial dragging NOR gate using two asymmetric stages. A free-space region between the nonlinear sections allows the pump and signal 2 to be brought into coincidence without nonlinear interaction.

gate of the same length); however, the pump is no longer standardized, but does not necessarily present any disadvantages from a systems point of view.

The studies of 1-D spatial soliton based logic gates of this section show that indefinitely cascadable, all-optical logic gates with fanouts of 2 or more can be obtained with large noise margin. The situation of ultimate interest for low-energy, ultrafast operation, is the use of fully-confined, spatio-temporal solitary waves. It is expected that the spatial collision and dragging interaction between these solitary waves will result in similar performance parameters to those obtained here. This is the subject of the next chapter of this thesis.

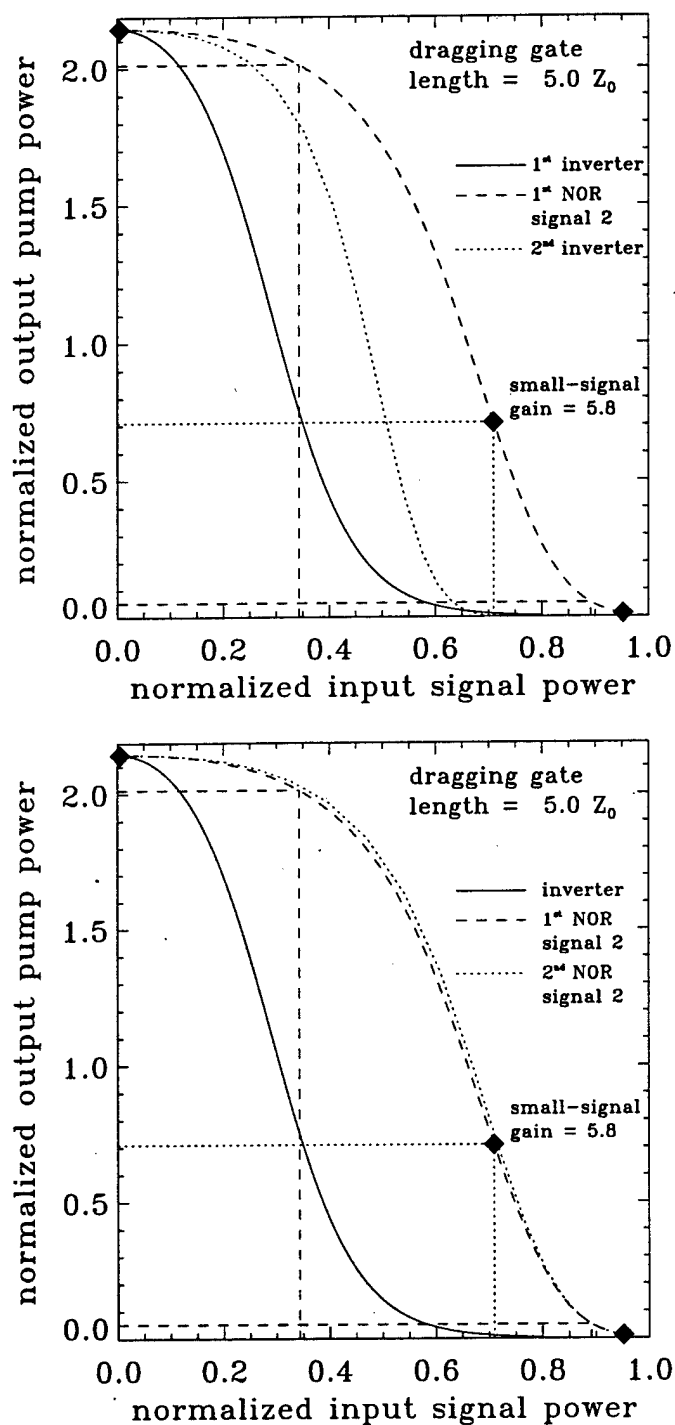


Figure 5.41: Transfer function for cascaded asymmetric NOR gate based on two $r = 3$ dragging gate stages of lengths $2 Z_0$ and $3 Z_0$ with normalized angle $\kappa = 0.8$. The signal 2 input of the first 2-NOR gate in the cascade is driven by an inverter with fanout 2.25, and the input to a cascaded inverter is driven by the 2-NOR output with fanout 3.0 (top) and the second signal input to the second 2-NOR gate is driven by the first with fanout 2.25 (bottom). The noise margins $NM_L = 0.34$ and $NM_H = 0.06$ for the 2-NOR gate, where the high noise margin represents 6.4% deviation about I_H , where the high noise margin represents 6.4% deviation about I_H .

Chapter 6

2-D Spatio-Temporal Solitary Wave Logic Gates

The ideal optical logic gate should be robust and operate with low energy at very high speed exceeding that obtainable by electronics. The first characteristic leads to the use of optical solitons which are robust carriers of information and are naturally binary due to their formation above a specified threshold [131]. In order to achieve low energy operation, an optical soliton logic gate should be based on one of the three geometries utilizing bright solitons that allow for complete three-dimensional confinement [2]: 1-D temporal solitons in fiber [59, 132], 2-D spatio-temporal solitary waves in slab waveguides [133, 134], and 3-D "light bullets" [2, 135, 136]. Finally, high-speed operation implies the use of ultra-fast, non-resonant nonlinearity, such as the third-order Kerr nonlinearity responsible for nonlinear refraction which is automatically phase-matched over broad angular and spectral bandwidths, allowing the use of solitons confined within the volume of approximately λ^3 [133]. Because of the relatively small size of the instantaneous Kerr nonlinearity, and the slow response time of large, resonant nonlinearities, simultaneous achievement of the second and third characteristics is generally considered difficult and leads to a fundamental "power delay" product [137]. Nevertheless, since nonlinear effects depend on peak intensity, low energy ultrafast switching in Kerr media may be possible by using tightly-focused beams of short duration. The limit to this scaling may be placed by the higher-order effects in the wave equation necessary to describe the propagation of optical radiation with large spatial and temporal bandwidth, as derived in Chapter 3, but there is reason to believe that balancing within the linear and nonlinear hierarchy of higher-order terms may be possible such that stabilized propagation results.

6.1 (2+1)-D Spatio-Temporal Propagation and Stability

Before logic gates can be studied, the stability of the pump solitary wave needs to be investigated. As mentioned previously, the angular deflection gates work on the principle that the pump does not broaden spatially, or at least broadens at an effective angle that is less than the induced deflection angle. It is also important for cascadability that the pump at the output is not significantly broader than the signal at the input since the gate length is based on the confocal parameter of the signal; however, imaging optics can be used to match the output pump width to the appropriate signal input width. The stability of the pump is determined by propagation using the (2+1)-D scalar spatio-temporal nonlinear wave equation (c.f. equation 2.99 or the scalar reduction of equation 3.170):

$$\begin{aligned}
 2ik_0 \frac{\partial A}{\partial z} + \frac{\partial^2 A}{\partial x^2} - k_0 k_0'' \frac{\partial^2 A}{\partial T^2} - \frac{i}{3} k_0 k_0''' \frac{\partial^3 A}{\partial T^3} - i \frac{k_0'}{k_0} \frac{\partial^3 A}{\partial T \partial x^2} \\
 + \frac{1}{12} k_0 k_0'''' \frac{\partial^4 A}{\partial T^4} + 2k_0^2 \frac{n_K}{n_0} |A|^2 A + 4ik_f n_K \left[\frac{n_0}{c} - \frac{k_0'}{2} \right] \frac{\partial |A|^2 A}{\partial T} \\
 + \frac{k_f}{2} \left\{ k_f + i \frac{2}{n_0} \left[\frac{n_0}{c} - \frac{k_0'}{2} \right] \frac{\partial}{\partial T} \right\} \int_0^\infty R_R(\tau) |A(T-\tau)|^2 A d\tau \\
 + 2k_0^2 \frac{n_4^{\text{eff}}}{n_0} |A|^4 A = 0,
 \end{aligned} \tag{6.1}$$

which is repeated here for convenience. An initial field envelope distribution is evolved by solving this equation using the split-step method as described in section 4.2.2.

It is expected that equation 6.1 does not possess stationary solutions because of the lack of compensation for the frequency down shifting caused by the Raman process, such as that provided by dispersive loss/gain [138]. Instead, the initial amplitude distributions in the simulations are the numerically-computed eigenmodes of the normalized (2+1)-D NLS equation with the total instantaneous cubic nonlinearity and the addition of the n_4 saturation term, as discussed in section 2.5. Note that there is reason to believe that approximate analytical solitary wave solutions to equation 6.1 exist in the absence of the Raman and quintic terms for restricted values of the material constants [139], for example when $\text{TOD} < 0$ (combined with space-time focusing) balances with shock [140]. In this thesis, the focus is instead on the ability of quintic saturation to maintain stability for more general situations.

Since the simulations are performed using real units and actual linear and nonlinear (except for the quintic nonlinear index) material parameters, the properties of fused silica, which is used as the example material in the simulations, are summarized. At $\lambda_f = 1.55 \mu\text{m}$, the material constants for fused silica are: $n_0 = 1.444$, $k'_0 = 4.88 \text{ fs}/\mu\text{m}$, $k''_0 = -0.0279 \text{ fs}^2/\mu\text{m}$, $k'''_0 = 0.151 \text{ fs}^3/\mu\text{m}$, $k''''_0 = -0.493 \text{ fs}^4/\mu\text{m}$, and $n_2^I = 3.3 \times 10^{-16} \text{ cm}^2/\text{W}$, from Appendix D. The linear properties were calculated using the Sellmeier coefficients of fused silica [30] ignoring the dispersive effects of the slab waveguide for simplicity; the full linear dispersion is used in the simulations. For these parameter values, $n_0/c - k'_0/2 = 2.38 \text{ fs}/\mu\text{m}$, meaning that shock will occur on the trailing edge of the solitary wave [30]. Fused silica has relatively small AGDD at $\lambda_f = 1.55 \mu\text{m}$, such that the scaling factor relating width to length $\sqrt{|k_0 k''_0|}/k'_0 = 0.08 \ll 1$. It is desirable to obtain a larger value so that the pulse duration will be longer for a given width (because of the scalings leading to equation 2.106, the pulse duration scales with the width as $\tau = w_0 \sqrt{|k_0 k''_0|}$) and to some extent reduce the effects of the higher-order, non-SVEA terms¹. In fact, it is because of this material scaling factor that the non-SVEA terms are more important than the non-paraxial terms. Larger AGDD can be obtained by properly tailoring waveguide dispersion [61] as recently demonstrated using an anti-resonant reflecting optical waveguide (ARROW) structure in AlGaAs for operation near $1.5 \mu\text{m}$ [37].

The parameters for the Raman response function were fit to values measured in fiber [141]: $R_0 = 2.63 \times 10^{-12} \text{ cm}^2/\text{W}\cdot\text{ps}^2$, $\gamma = 65.6 \text{ rad/ps}$, and $\Omega_f = 89.0 \text{ rad/ps}$, leading to a peak in the Raman gain at 13.2 THz , as discussed in Appendix B. For propagation of the $\sigma = 0.5$ eigenmode of the cubic-quintic NLS equation with $w_0 = 39.6 \mu\text{m}$, $n_4^{\text{eff}} = -0.5n_2/I_0 = -2.5 \times 10^{-28} \text{ cm}^4/\text{W}^2 \sim -2300 n_2^I$, where $I_0 = \epsilon_0 c n_0 |A_0|^2/2$ is the peak intensity. The value of n_4^{eff} that might be obtainable in fused silica can be estimated. Assuming that the third-harmonic cascaded contribution dominates the total intrinsic contribution given by $n_4^I - n_2^I/2n_0$ and that $n_2 \sim 3\chi_{xxxx}/8n_0$, the quintic index from equation 3.167 can be written

$$n_4^{\text{eff}} \approx \frac{2}{3} \frac{n_0 n_2^I}{n_2^I - n^2(3\omega_0)} \approx \frac{1}{3} \frac{n_2^I}{n_0 - n(3\omega_0)}. \quad (6.2)$$

Using the Sellmeier data, the refractive index at the third-harmonic wavelength $\lambda_f/3 = 0.517 \mu\text{m}$ is $n(3\omega_0) = 1.461$, resulting in $n_4^{\text{eff}} \approx -2.1 \times 10^{-30} \text{ cm}^4/\text{W}^2 = -20n_2^I$, a factor of about 120 smaller than the value needed in the simulations to obtain $\sigma = 0.5$ for the present choice of the width parameter w_0 . This simple estimate indicates that the value of n_4^{eff} necessary for stable solitary-wave propagation (as shown later) may not be obtainable in bulk fused silica; however, utilization of quasi-phase matching or three-photon resonance enhancement could substantially enhance the cascaded contribution. For example, a long period ($\sim 23.5 \mu\text{m}$) index grating would provide the necessary phase matching to enhance n_4^{casc} by the requisite 2 orders of magnitude, but additional analysis may be necessary since this approaches the resonance condition for which equation 3.137 for the generation of the third-harmonic is not valid. In the case of resonant enhancement, the detrimental effects of three-photon absorption [142] on solitary wave propagation would need to be considered. Alternatively, the value of n_4^{eff} needed in the simulations can be reduced (because $I_0 \propto 1/w_0^2$) by making w_0 smaller than the value of $39.6 \mu\text{m}$ used here (while maintaining $\sigma = 0.5$), or by reducing σ itself, resulting in more broadening of the spatio-temporal wave. Decreasing w_0 to $3.6 \mu\text{m}$ with $\sigma = 0.5$ to obtain n_4^{eff} compatible with the estimate for fused silica results in a spatial FWHM of $\sim 3.7 \mu\text{m}$ and temporal FWHM of $\sim 1.5 \text{ fs}$ so that the temporal duration would be less than an optical cycle which is clearly outside the regime of validity of the derivation in Chapter 3. In combination, grating-assisted phase-matching with smaller w_0 could result in achieving the appropriate saturation condition without violating the assumptions under which the derivation of section 3 was performed. For comparison that the quintic index for PTS at $\lambda_f = 1.6 \mu\text{m}$ is $n_4^{\text{eff}} = -8.5 \times 10^{-22} \text{ cm}^4/\text{W}^2 \sim -176 n_2^I$ [125], and for AlGaAs at sub-half bandgap $n_4^{\text{eff}} = -5.0 \times 10^{-23} \text{ cm}^4/\text{W}^2 \sim -3500 n_2^I$ (due to three-photon resonant enhancement [143]), which is more consistent with the simulation parameters.

In order to understand the simulation results of this section, a brief summary of the results of section 2.4.4 on the effects of the higher-order terms is presented. Positive third-order dispersion (TOD) produces oscillatory behavior on the temporal trailing edge, with slight delay of the peak in the reduced-time coordinate frame. When combined with AGDD, this oscillation dampens and the peak advances [30] instead of delays. In the soliton regime when AGDD is balanced by self-phase modulation (SPM), positive TOD again causes delay of the peak with narrowing of the temporal profile. Fourth-order dispersion

¹However, increasing the temporal duration with fixed spatial width also leads to increased solitary wave energy.

of the same sign as group-delay dispersion (as it is here) will enhance the effect of GDD in the spectral wings, while the opposite sign would reduce the effect. Space-time focusing is the off-axis projection of the group velocity onto the direction of propagation [144] and describes the natural curvature of the energy front due to spatio-temporal diffraction. The shock term manifests itself as an intensity-dependent group velocity [30], with more intense portions of the envelope traveling slower than the less intense portions resulting in a steepening of the trailing edge. The shock is dissipated by dispersion because of spectral broadening occurring at the trailing edge. Raman scattering results in a continuous frequency downshift² by providing gain at lower frequencies at the expense of higher frequencies. This effect is referred to as the soliton self-frequency shift [145]. In the AGDD regime, the downshifted frequencies travel with a slower group velocity thereby delaying the wave in the reduced time coordinates. For the chosen parameters, both higher-order nonlinear temporal terms contribute to delay relative to the mean group velocity $v_g = 1/k'_0$. A separate phenomenon not directly associated with the higher-order terms, but nevertheless responsible for important effects, is spatio-temporal coupling [146]. Spatio-temporal coupling is a property of multi-dimensional NLS propagation that links width and duration, such that a change in one produces a change in the other. In the AGDD and self-focusing regime, the width and duration can evolve in-phase.

Spatio-temporal propagation without quintic saturation is presented first. The results obtained are more indicative of materials such as fused silica where the quintic nonlinearity may not play much of a role for typical peak intensity levels. In the subsequent section, the results for propagation with the saturation term are presented with further discussions on the implications for using spatio-temporal solitary waves in optical switching. Materials for which these results might be indicative are AlGaAs [143] and PTS [147] due to their large, negative values of n_4 .

For all simulations, $w_0 = 39.6 \mu\text{m}$ so that the unsaturated $\sigma = 0.0$ eigenmode with peak normalized amplitude $U_0 = 1.94$ and the saturated $\sigma = 0.5$ eigenmode with $U_0 = 1.0$, which would be used as pump spatio-temporal waves, have initial intensity full-widths at half-maximum (FWHM) spatial widths of $40.7 \mu\text{m}$ and corresponding temporal FWHM durations of 16.5 fs. The resulting total energies are about 6.0 nJ for the $\sigma = 0$ wave and 11 nJ for the $\sigma = 0.5$ wave, if confined by a $2 \mu\text{m}$ thick waveguide; these energies can be reduced by three orders-of-magnitude or more by using enhanced nonlinear materials such as AlGaAs [143] combined with tailored dispersion to achieve AGDD [37]. The propagation distance for the stability studies is set at $15 Z_p$, where Z_p is the confocal distance which is calculated numerically as twice the distance over which the initially transform-limited intensity profile broadens by a factor of $\sqrt{2}$ in linear propagation. The simulations are performed with a 1024 by 1024 computational grid which extends from $-1020 \mu\text{m}$ to $1020 \mu\text{m}$ in the spatial dimension and from -0.624 ps to 0.623 ps in the temporal dimension. In the spatio-temporal frequency domain, the grid extends from $-1.57 \text{ rad}/\mu\text{m}$ to $1.57 \text{ rad}/\mu\text{m}$ in the spatial frequency dimension and from -2570 rad/ps (-409 THz) to 2580 rad/ps (411 THz) in the temporal frequency dimension. The grids in both domains are sub-sampled to 512 by 512 for the plots.

6.1.1 Propagation without Saturation

For comparison purposes, the linear propagation of the $\sigma = 0$ eigenmode, obtained from the (2+1)-D cubic NLS equation (without saturation) is first examined. Under the paraxial and SVE approximations, the spatio-temporal wave broadens spatially from $40.7 \mu\text{m}$ to $431 \mu\text{m}$ (a factor of 10.6) and temporally from 16.5 fs to 174 fs (a factor of 10.5) after propagation over $15 Z_p$. Here, the confocal distance $Z_p = 0.451 k_0 w_0^2 = 4.17 \text{ mm}$. It is important to note that without the higher-order linear terms, the energy front (defined here as the position of the temporal centroid across the spatial profile) is stationary in the reduced time coordinate frame, as previously shown in Figure 4.5 over $7.5 Z_p$. Under SVEA, the energy front lies on a plane perpendicular to the direction of propagation, while the phase front has quadratic curvature. The non-SVEA space-time focusing term adds quadratic curvature to the energy front as well, while the addition of successive non-paraxial and non-SVEA non-paraxial terms (i.e. the space-time dispersion terms) correct the phase and energy fronts to lie on the perimeter of a circle instead of a parabola.

Figure 4.6 shows intensity contour plots of the $\sigma = 0$ eigenmode in linear propagation over $7.5 Z_p$ using the full spatio-temporal phase given by equation 4.12. More detail about the spatio-temporal wave after linear propagation over $15 Z_p$ is shown in Figure 6.1. The dashed curve in the top figure denotes the position of the energy front across the spatial profile, which is indicated with the temporal centroid. An alternate, but numerically less robust, measure is the position of the peak of each temporal slice. After propagation, the solitary wave broadens by a factor of 10.5 (to $430 \mu\text{m}$) and 10.3 (to 170 fs) in the spatial and temporal dimensions, respectively. The top and middle plots that the energy front is no longer flat. Due to the interaction between AGDD and positive TOD, bunching at the leading edge and tailing at the trailing edge is clearly seen in the top plot. As a result, the peak of the temporal profile at the center of the wave is advanced (by 16.3 fs) in reduced-time coordinates (while the temporal centroid is delayed by 7.6 fs).

The middle figure plots the energy front curve on a smaller scale, which indicates that the temporal positions in the wings are delayed relative to the center. This effect is the result of the curvature of the energy front due to space-time focusing, such

²Continuous downshift only occurs for pulses that are not strongly chirped where the interacting frequencies overlap at the same positions under the envelope, which occurs for a temporal soliton; otherwise, causality would be violated.

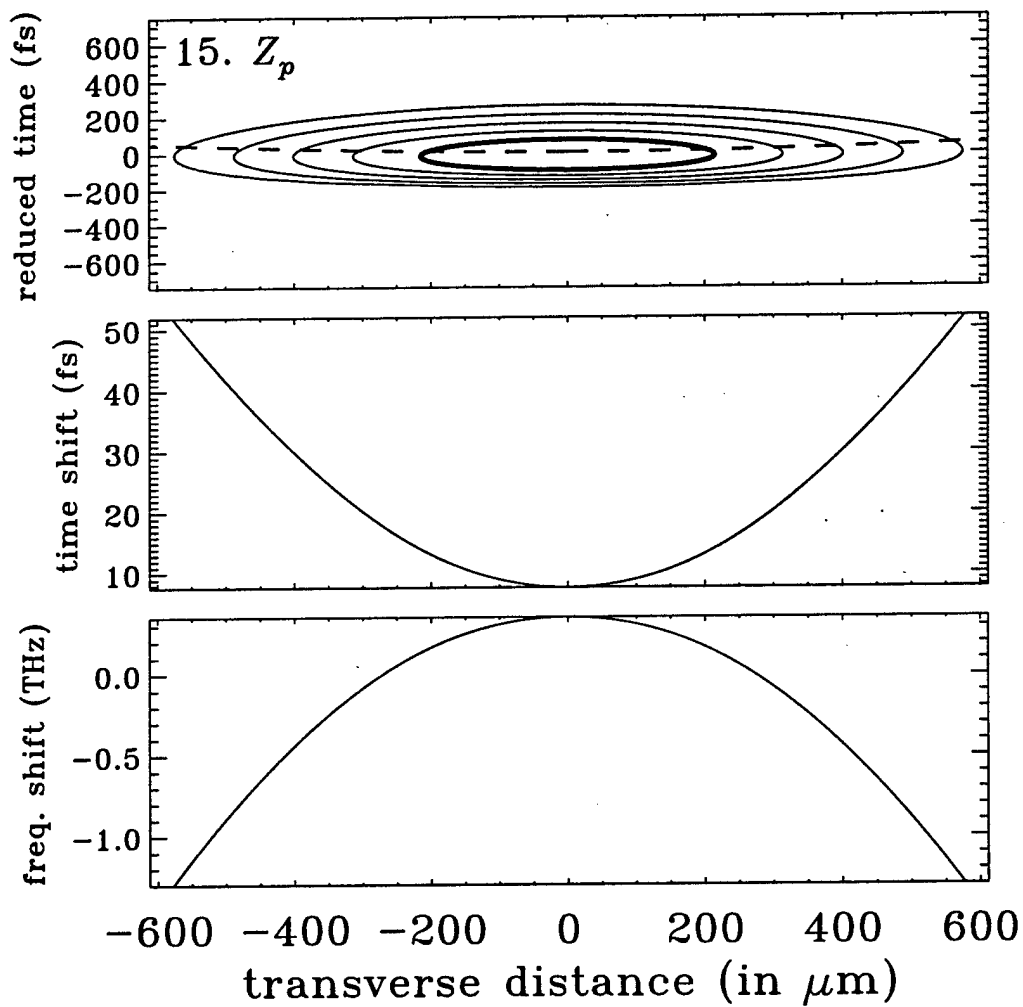


Figure 6.1: Higher-order linear propagation of the $\sigma = 0$ eigenmode, of spatial FWHM $40.7 \mu\text{m}$ and temporal FWHM 16.5 fs. The dashed curve in the top frame marks the position of the temporal centroid across the spatial profile, which broadens to $430 \mu\text{m}$ FWHM. At $x = 0$, the central portion is delayed by 7.63 fs. The middle frame plots the temporal centroid on a smaller time scale and reveals a delay in the wings of 44.3 fs relative to the center. The bottom frame plots the frequency shift centroid across the spatial profile indicating that the wings are downshifted by 1.65 THz relative to the center.

that projection of the group velocity onto the axis of propagation becomes smaller for increasingly off-axis rays. Because of the temporal variation across the wave profile, there must be a corresponding local frequency shift. This is shown in the bottom figure, which plots the centroid frequency of the power spectrum in each temporal slice across the wave. The central portion of the wave is upshifted, which corresponds to advancement in the AGDD regime (and occurs for the temporal peak at the center, not the temporal centroid), while the wings are downshifted, which corresponds to delay. Diffraction of lower frequency components is greater than for higher frequency components, thus lower frequencies dominate the wings and high frequencies dominate the central portion. Note that, even though the calculation shown a nonzero frequency shift across the wave profile, the overall spatio-temporal power spectrum cannot change in linear propagation. The spectral phase is modified along temporal slices parallel (and spatial slices perpendicular) to the direction of propagation however, meaning that the wave is no longer transform limited.

The next question to address is whether a spatio-temporal solitary wave results under nonlinear propagation without quintic saturation. Figure 6.2 shows the result of this propagation. This simulation verifies that the higher-order terms (without saturation) prevent the collapse that would occur in multi-dimensional NLS propagation in the AGDD regime, but the simulation also shows that these terms do not stabilize, resulting instead in weak diffraction/dispersion. The broadening factors in this case are reduced considerably to 3.96 and 4.12, respectively, over linear propagation, also indicating that strict spatio-temporal coupling does not occur due to the symmetry breaking of the higher-order temporal terms. Collapse is prevented because the wave is initially strongly perturbed by the higher-order (mainly Raman) terms, but without saturation as a stabilization mechanism to oppose these perturbations, the tightly localized wave begins to spread.

The dashed curves in Figure 6.2 indicate the temporal position of the energy front across the spatial profile. The central portion of the wave is delayed 35.5 fs (temporal centroid), with a spectral downshift of 2.83 THz. The delay (and therefore downshift) is largely a result of the Raman and shock terms, but since AGDD and SPM approximately balance, it is expected that positive TOD is also a contributing factor. Since some of this delay is caused by nonlinear effects, the central portions of the wave are more strongly affected than the wings, thereby partially counteracting space-time focusing. The result is that the relative time delay of 19.3 fs between the wings and center is reduced over the relative delay of 44.3 fs of the linear case. Like space-time focusing, these are true spatio-temporal effects and have no analogs in purely temporal propagation, but it is clear that the linear propagation effects dominate. Now, because of the higher-order nonlinear terms (and the balancing of AGDD with SPM), the central portion is spectrally downshifted, and the difference in the downshift between the wings and center is 1.17 THz, rather than 1.65 THz as for linear propagation.

The overall temporal centroid of the spatio-temporal wave is delayed by 38.8 fs relative to the center of the reduced-time coordinate system. For this propagation distance, the time shift is not resolvable and therefore considered unimportant for optical switching applications. For cascaded logic gates, though, this shift can be important if the temporal broadening is slight, which is not the case here. Figure 6.3 plots the initial and final spatio-temporal frequency power spectrum, showing downshift and substantial reduction of both the spatial and temporal frequency content. The change in the power spectrum is caused by the nonlinear terms (since the linear terms only modify spectral phase), and the cause of the spectral reduction is discussed in the section on propagation with saturation. The overall temporal frequency downshift of 3.05 THz is not spectrally-resolvable and also has negligible consequences for optical switching.

A plot of the evolution of the spatial and temporal FWHM versus propagation distance is shown in Figure 6.4 for the cases with and without the Raman term, along with a comparison to linear propagation. In the absence of Raman scattering, the solitary wave begins to collapse, but this collapse is eventually arrested by the remaining higher-order terms. The minimum size of 21.4 μm and 7.57 fs is reached at a distance of 14.6 Z_p . During collapse, the effects of the higher-order linear terms and nonlinear SPM and optical shock become more important and their effective interaction lengths increase due to spatio-temporal confinement. As a result, these terms induce a larger (than with Raman) overall temporal delay of 47.8 fs, but slightly smaller overall temporal frequency downshift of 2.67 THz. Spectral broadening via SPM (i.e. four-wave mixing) is enhanced during the collapse process (the threshold for continuum generation is about the critical power for self-focusing [148]) such that the temporal frequency spectrum is broadened by a factor 1.17, versus narrowing by a factor 0.641 as in the case with Raman scattering. Figure 6.4 also shows that the presence of Raman scattering immediately prevents collapse, resulting in broadening of the solitary wave, thereby reducing the interaction length and the effects of the other higher-order terms.

For optical switching applications, a "solitary wave" need only be stable (i.e. maintain approximate width) over the length of the logic gate, which may be as short as 2 Z_0 for a properly optimized geometry as shown in the previous chapter, where Z_0 is the confocal distance of the initial signal, and $Z_0 = 3.21 Z_p$. It is clear from Figure 6.4 that non-saturated propagation without Raman scattering satisfies this criterion over the entire distance of the simulation. Unfortunately, Raman scattering is unavoidable in fused silica (and essentially all media except for monatomic vapors) such that the wave does not maintain a nearly stable form without saturation. Nevertheless, non-saturated propagation even with Raman scattering may sufficiently satisfy the criterion over short gate lengths as to be useful in the angular deflection switching geometries in which some amount of spatial broadening can be tolerated.

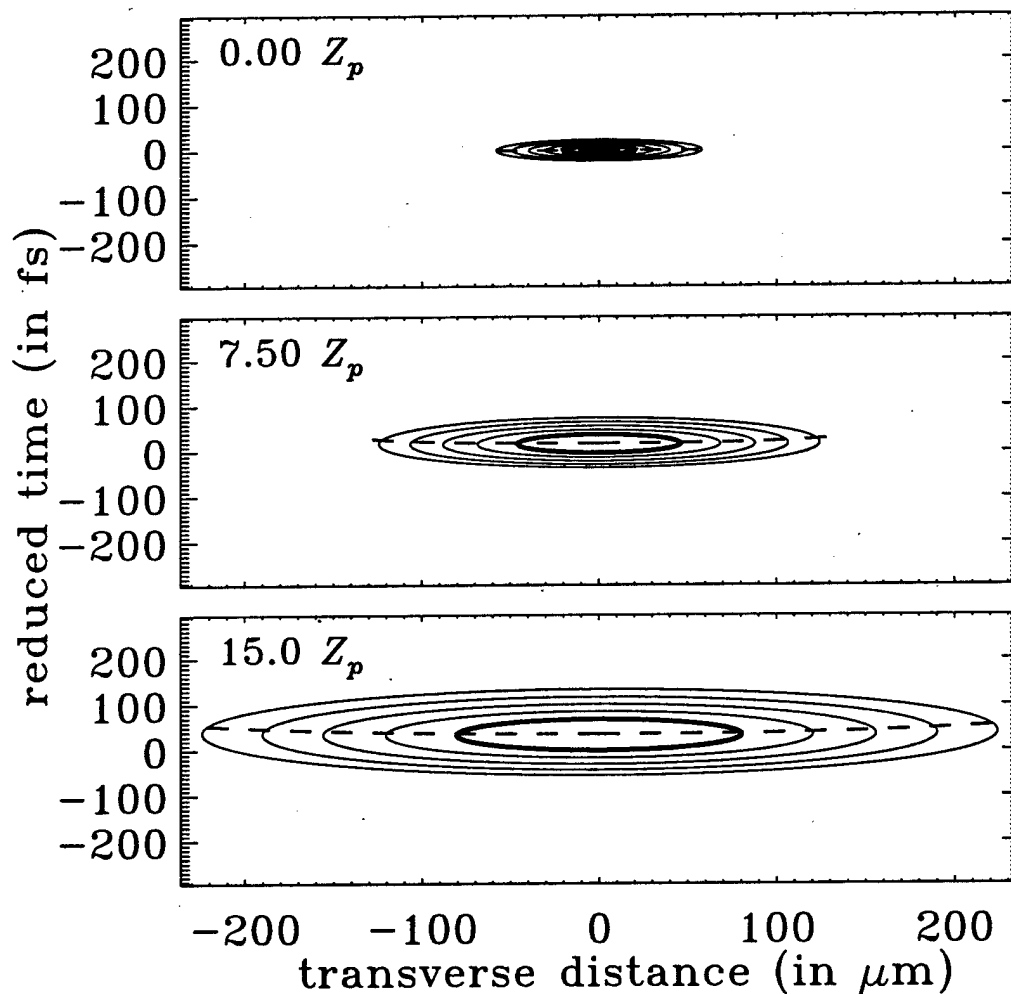


Figure 6.2: Propagation of the $\sigma = 0$ eigenmode including all of the higher-order terms except saturation. The spatio-temporal wave of initial transverse FWHM $40.7 \mu\text{m}$ and temporal FWHM 16.5 fs (top frame) broadens during propagation to $91.9 \mu\text{m}$ and 40.6 fs at $7.5 Z_p$ (middle frame) and $161 \mu\text{m}$ and 67.9 fs at $15 Z_p$ (bottom frame). The dashed curve in each frame marks the position of the temporal centroid across the spatial profile. At the end of propagation, the central portion of the wave (at $x = 0$) is delayed by 35.5 fs , and the wings are delayed by 19.3 fs relative to center. The contours are at -3 dB intervals relative to the peak intensity in each frame.

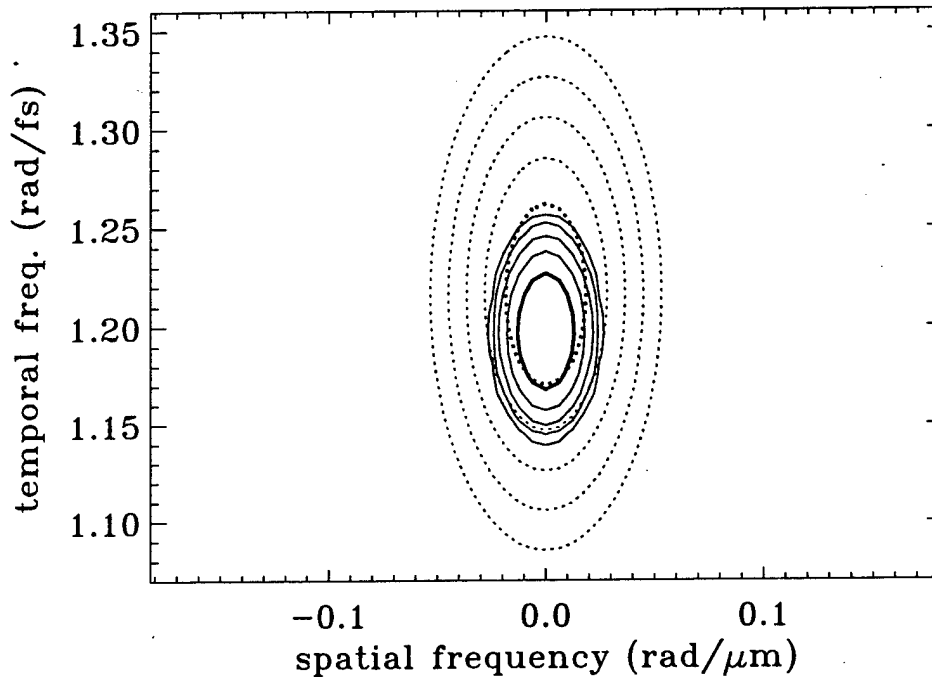


Figure 6.3: Spatio-temporal frequency profile of the $\sigma = 0$ eigenmode initially (dotted contours) and after higher-order propagation over $15 Z_p$ (solid contours). The initial spectral FWHM $0.0371 \text{ rad}/\mu\text{m}$ (0.363 deg.) and 92.0 rad/ps (14.6 THz) decreases to $0.0261 \text{ rad}/\mu\text{m}$ (0.255 deg.) and 58.9 rad/ps (9.38 THz). The overall frequency centroid is downshifted by 19.1 rad/ps (3.05 THz). The contours are at -3 dB intervals relative to the respective peaks.

6.1.2 Propagation with Saturation

Stable propagation of the $\sigma = 0.5$ eigenmode, for which the peak intensity is one-half the "saturation" intensity, was verified in Figure 2.20 with the addition of just the quintic saturation term to the multi-dimensional NLS equation. Because the spatio-temporal wave is unchanged, it is designated as being a true solitary wave. The same stabilized solitary-wave behavior occurs in the fully 3-D case as well [2]. The confocal distance for this spatio-temporal wave $Z_p = 0.559 k_0 w_0^2 = 5.14 \text{ mm}$. Due to the difference in shape, the $\sigma = 0.5$ eigenmode has a larger constant factor, and therefore, a slightly longer confocal distance than the $\sigma = 0$ eigenmode. The main goal of this section is to determine if stability is maintained under the influence of the remaining higher-order terms.

The propagation of the $\sigma = 0.5$ eigenmode with all of the higher-order terms is shown in Figure 6.5. It is clearly seen that spatio-temporal broadening occurs, but not to the extent that it occurs for propagation without saturation. Here, the broadening factors are just 1.46 and 1.81 respectively, with more broadening occurring in the temporal dimension because the relevant higher-order terms, except for saturation, depend on time, and because the downshift causes a slight increase in the effective AGDD at the new central frequency due to TOD and FOD, leading to a larger width to length scaling factor of about 0.1. Even though both the $\sigma = 0$ and $\sigma = 0.5$ spatio-temporal waves have the same initial FWHM, the saturated eigenmode is more strongly affected by the higher-order nonlinear terms because the initial peak intensity is more than a factor of 2 greater. In addition, since saturation tends to stabilize propagation so as to maintain width, duration, and peak intensity, the spatio-temporal wave experiences the higher-order effects over much longer distance than for propagation without saturation.

As far as optical switching is concerned, there are two main effects that arise due to this enhanced interaction with the higher-order terms that need to be considered. The first is that the solitary wave in Figure 6.5 is delayed overall by 127 fs, or 7.7 times the initial duration. In cascaded switching operation, this delay can be compensated for by retiming subsequent inputs, but may affect operation if the solitary-wave amplitude is subject to fluctuation which would result in resolvable pulse-to-pulse delay variation at the input of the next gate. The second effect is the spectral downshift, as shown in Figure 6.6. The overall downshift of 10.5 THz is slightly less than resolvable, but may also lead to difficulties in cascaded operation. The difference in central frequency of two interacting solitary waves leads to a difference in group-velocity, which, in the absence

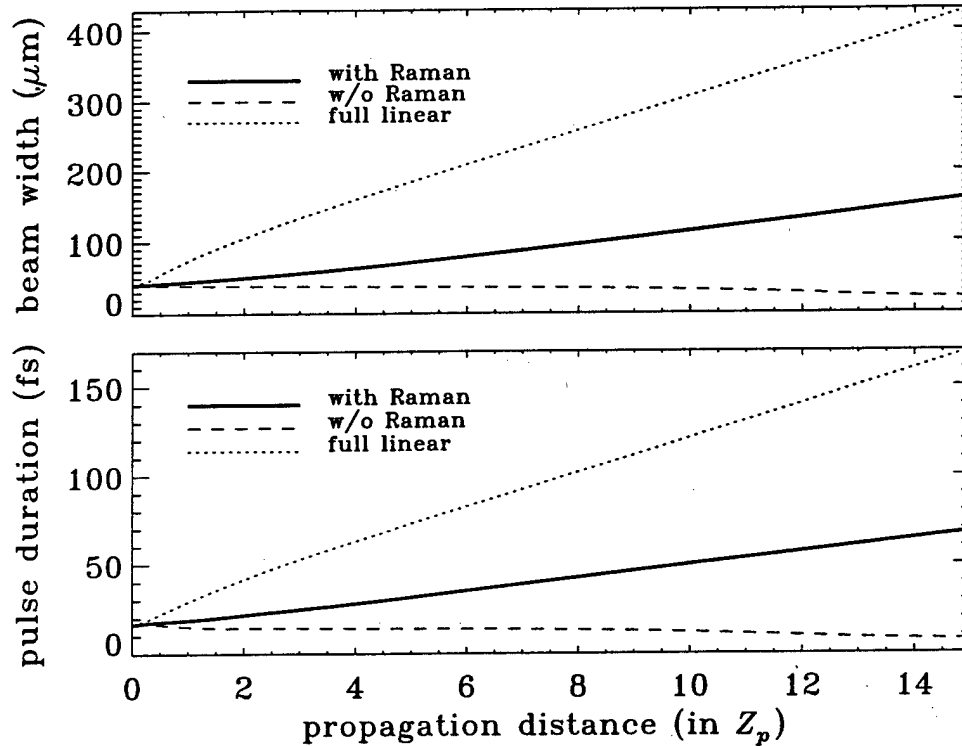


Figure 6.4: Evolution of spatial and temporal FWHM with propagation distance for the $\sigma = 0$ cubic eigenmode without saturation. The solid curves show the spatial (top) and temporal (bottom) FWHM for propagation with Raman scattering, the dashed curves show the spatial (top) and temporal (bottom) FWHM without Raman scattering, while the dotted curves show the same for purely linear higher-order propagation.

of temporal trapping, will reduce the interaction length and may result in lower contrast operation. Extensive study of the temporal trapping phenomenon has been performed for soliton interaction in fiber, where the difference in group velocity is due to birefringent [57,59] or dispersive [63,64] walkoff. To date, no such study has been performed in the multi-dimensional, spatio-temporal case, but it is expected that there is a similar maximum difference in group-velocity for which temporal trapping will result. Compensation for this dispersive walkoff based on waveguide dichroism will be used in section 6.3 for cascaded operation.

As for the simulations of the previous section, other spatio-temporal effects occur that may not affect optical switching, but are nonetheless worth discussion. As shown by the dashed curves in Figure 6.5, the curvature of the solitary wave energy front is opposite to that which would be caused by space-time focusing alone; now, the wings are *advanced* by 18.4 fs relative to the center. This is a result of the enhanced action of the nonlinear Raman and shock terms, which delay the center portion more than the wings and now overtake space-time focusing. The steepening of the *trailing* edge is clearly seen by the bunching of contour lines. In addition to the shock term producing an intensity-dependent group delay resulting in self-steepening, the Raman-induced self-frequency shift, also being intensity dependent, produces an additional self-steepening effect since the central portion will be downshifted the most and hence, due to AGDD, delayed more in the coordinate frame than the wings.

Figure 6.6 shows that the spatio-temporal frequency spectrum is reduced by a factor of 0.475 and 0.470 in the spatial and temporal frequency dimensions, respectively. Narrowing of the temporal frequency spectrum is a manifestation of the Raman gain process, which does not generate new frequencies, but amplifies existing red frequencies at the expense of existing blue frequencies, and results in temporal broadening of the solitary-wave profile. Due to space-time coupling [146], the spatial profile must broaden as well without significant phase curvature in order to maintain stabilized, nearly transform-limited, solitary-wave propagation. The continuous nature of the spectral downshifting is a result of four-wave mixing which constantly creates new frequencies, and can lead to downshifting by an amount greater than allowed by the initial spectral content, but this effect is slight due to the approximate balancing between SPM and GDD. Over the propagation distance of $15 Z_p$, an initial spectral downshift of 10.5 THz would correspond to a time delay of 144 fs. Since the actual overall time delay is 127 fs, this simple analysis indicates that most of the downshift occurs over the first few Z_p of propagation.

Performing the same simulation without the Raman term reveals that the solitary wave is nearly unchanged after propagating

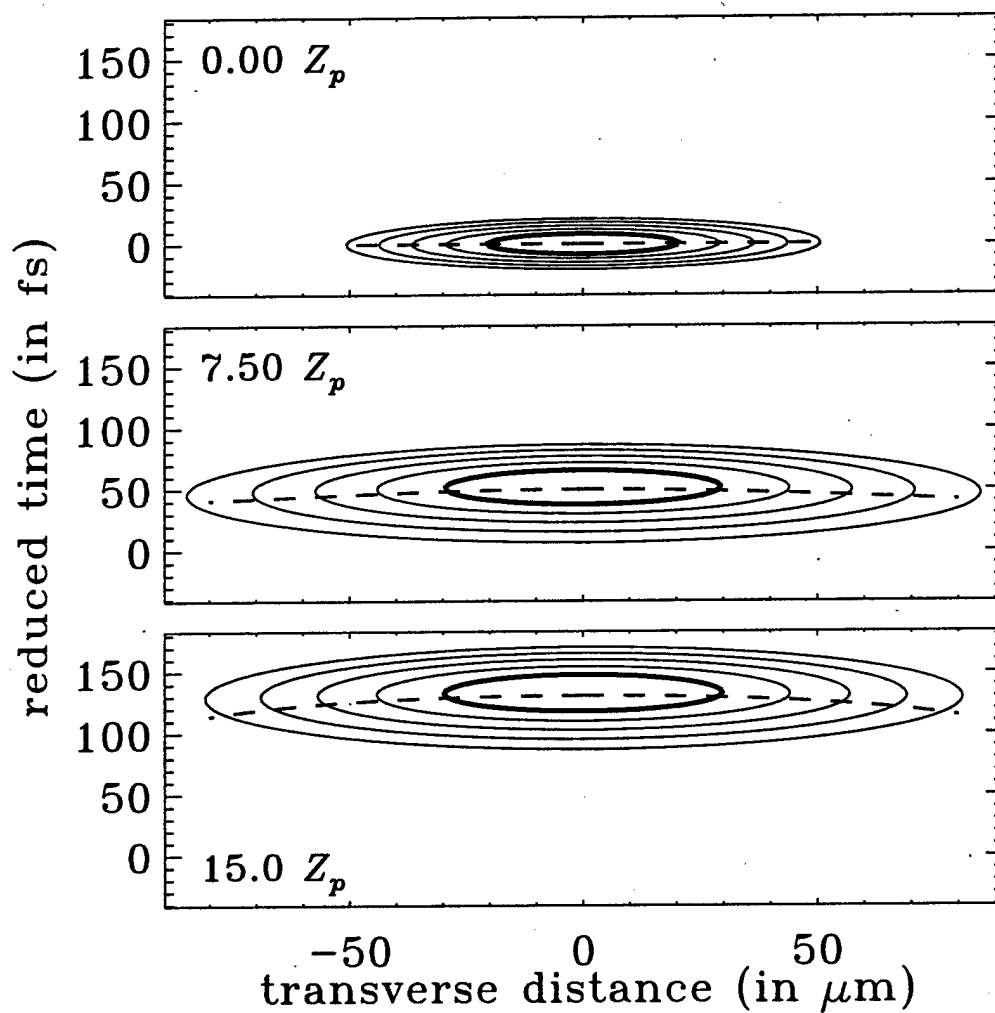


Figure 6.5: Stabilized propagation of the $\sigma = 0.5$ eigenmode including all of the higher-order terms. The spatio-temporal wave of initial transverse FWHM $40.7 \mu\text{m}$ and temporal FWHM 16.5 fs (top frame) broadens during propagation to $58.9 \mu\text{m}$ and 27.9 fs at $7.5 Z_p$ (middle frame) and $59.5 \mu\text{m}$ and 29.8 fs at $15 Z_p$ (bottom frame). The dashed curve in each frame marks the position of the temporal centroid across the spatial profile. At the end of propagation, the central portion of the wave (at $x = 0$) is delayed by 131 fs , and the wings are advanced by 18.4 fs relative to center. The contours are at -3 dB intervals relative to the peak intensity in each frame.

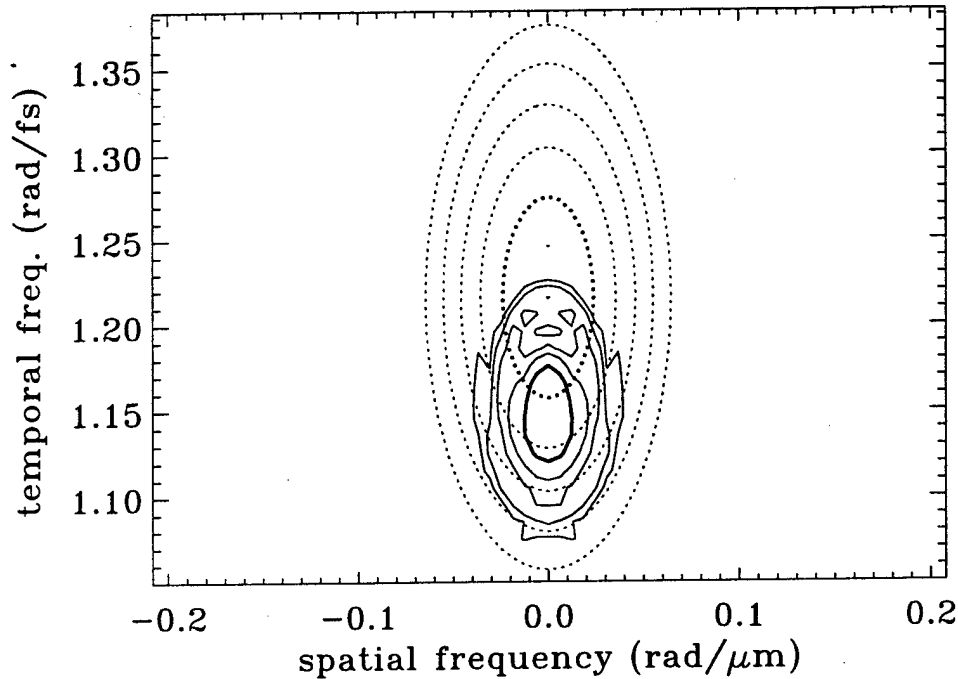


Figure 6.6: Spatio-temporal frequency profile of the $\sigma = 0.5$ eigenmode initially (dotted contours) and at $z = 15 Z_p$ (solid contours). The spatial frequency FWHM narrows from $0.0478 \text{ rad}/\mu\text{m}$ (0.467 deg.) to $0.0227 \text{ rad}/\mu\text{m}$ (0.222 deg.) and the temporal frequency FWHM narrows from 118 rad/ps (18.8 THz) to 55.5 rad/ps (8.84 THz). The overall centroid of the frequency spectrum is downshifted by 10.5 THz and oscillatory structure develops on the blue side. The contours are at -3 dB intervals relative to the respective peaks.

$15 Z_p$, with final spatial FWHM $40.7 \mu\text{m}$ and temporal FWHM 16.1 fs , with overall time delay of 50.4 fs . The evolution of the FWHM versus propagation distance for the saturated eigenmode with and without Raman scattering is shown in Figure 6.7. These curves show that, without the Raman term, the solitary wave does not change significantly during propagation, and that with the Raman term, most of the broadening occurs in about the first $6 Z_p$ of propagation, with nearly stable propagation occurring afterwards, where the solitary wave settles into a form that is much less strongly affected by Raman scattering. If this form were used as the initial condition instead of the $\sigma = 0.5$ eigenmode solution to the cubic-quintic NLS equation, then much less broadening would occur over the same propagation distance. Nevertheless, similar to the case without saturation, Raman scattering is the most detrimental effect to solitary wave propagation, but, in this case, saturation effectively stabilizes against the effects of the remaining higher-order terms. It is important to keep in mind that, even in the presence of all higher-order effects, the solitary wave does not diffract/disperse as quickly as propagation without saturation. Over a distance of $5 Z_p$, the spatio-temporal wave without quintic saturation would spatially broaden by a factor of 1.77 , while with saturation, the broadening factor is reduced considerably to just 1.39 . The important point here is that for typical switching geometries based on spatial discrimination of the output state, the smaller the spatial broadening factor, the higher the contrast of the switching operation. Therefore, while quintic saturation (or some other stabilization mechanism) may not be absolutely necessary, it is clearly desirable for all-optical switching devices based upon spatio-temporal solitary waves.

The results of the simulations show that without quintic saturation, due to spatio-temporal broadening instigated mainly by Raman scattering, nonlinear spatio-temporal waves may be restricted in optical switching applications to gates of short lengths, while, with the stabilization provided by quintic saturation, the resultant solitary waves should be useful in much more general switching geometries in which the output of one switching stage is used as a control input to a subsequent stage. While spatial broadening is detrimental to solitary-wave based switching devices, the corresponding lengthening of the temporal duration can be beneficial by increasing the timing tolerance [3], thereby relaxing synchronization issues.

Finally, the results of this section lend insight into additional stabilization mechanisms, such as the balancing between the linear effects of negative TOD and space-time focusing, and the nonlinear effect of optical shock. Here, with AGDD balancing SPM, the linear terms will act to advance the center and delay the wings, while shock will predominantly delay the center,

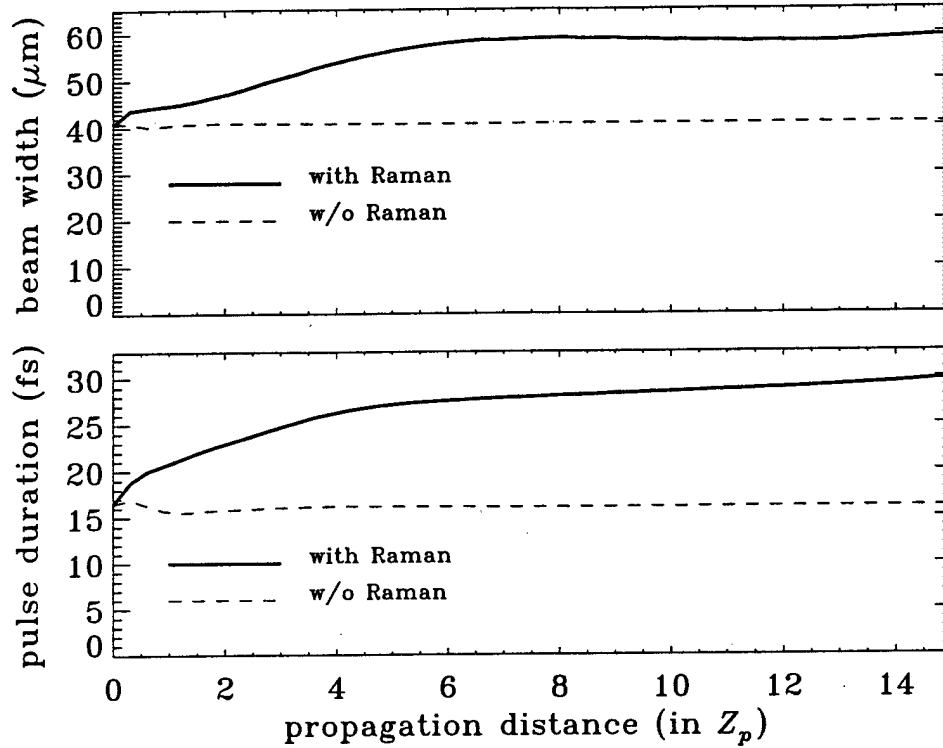


Figure 6.7: Evolution of spatial and temporal FWHM with propagation distance for the saturated $\sigma = 0.5$ cubic-quintic eigenmode. The solid curves show the spatial (top) and temporal (bottom) FWHM for propagation with Raman scattering, and the dashed curves show the spatial (top) and temporal (bottom) FWHM without Raman scattering.

resulting in a flat energy front. This, along with the balancing between the non-paraxial linear and nonlinear terms obtained in equation 3.161 to flatten higher-order curvature of the phase front, may also explain the recent results based on direct simulation of Maxwell's equations [133] which demonstrated stabilized propagation of solitary waves with large angular (non-paraxial) and temporal bandwidths (non-slowly varying envelope), in the absence of the Raman and quintic nonlinearities.

6.2 2-D Spatio-Temporal Dragging Logic Gates

This section considers the vectorial interaction between two solitary waves for use in single-stage optical logic. The results of this section will identify appropriate operating parameters for use in the inverter and 2-NOR cascading studies in section 6.3. Based on the results of Chapter 5 and the long simulation times in (2+1)-D, only the dragging interaction is considered, for which the best performance is expected. The geometry of this spatial dragging interaction between spatio-temporal solitary waves is illustrated in Figure 1.20. In addition, because of the long gate lengths (which are about 4 to 5 times longer than the 1-D soliton gates), which result from the large spatial width to length scaling imposed by the small AGDD coefficient of fused silica, linear absorption is neglected. Note, however, that the product of peak intensity and total propagation distance is small enough that two-photon absorption becomes a negligible effect according to equation 5.40, which serves as a good estimate for the 2-D case. The gate lengths can be reduced considerably either by using materials with larger AGDD coefficients or by tailoring the waveguide dispersion properties for which the spatial width can be decreased without affecting temporal duration, in which case linear absorption would no longer be a dominant effect.

The extension of the scalar equation to include nonlinear coupling with the orthogonal linear polarization results in (c.f. equation 4.47)

$$\begin{aligned}
 2ik_0 \frac{\partial A_x}{\partial z} + \frac{\partial^2 A_x}{\partial x^2} - k_0 k_0'' \frac{\partial^2 A_x}{\partial T^2} - \frac{i}{3} k_0 k_0''' \frac{\partial^3 A_x}{\partial T^3} - i \frac{k_0'}{k_0} \frac{\partial^3 A_x}{\partial T \partial x^2} \\
 + \frac{1}{12} k_0 k_0'''' \frac{\partial^4 A_x}{\partial T^4} + 2k_0^2 \frac{n_K}{n_0} \left[|A_x|^2 + 2\Delta_K |A_y|^2 \right] A_x
 \end{aligned} \tag{6.3}$$

$$\begin{aligned}
& + 4ik_f n_K \left[\frac{n_0}{c} - \frac{k'_0}{2} \right] \frac{\partial}{\partial T} \left[|A_x|^2 A_x + 2\Delta_K |A_y|^2 A_x \right] \\
& + \frac{k_f}{2} \left\{ k_f + i \frac{2}{n_0} \left[\frac{n_0}{c} - \frac{k'_0}{2} \right] \frac{\partial}{\partial T} \right\} \left[\int_0^\infty R_R(\tau) |A_x(T-\tau)|^2 A_x d\tau \right. \\
& \quad + \Delta_R \int_0^\infty R_R(\tau) |A_y(T-\tau)|^2 A_x d\tau \\
& \quad \left. + \gamma_R \int_0^\infty R_R(\tau) A_y^*(T-\tau) A_x(T-\tau) A_y d\tau \right] \\
& + 2k_0^2 \frac{n_4^{\text{eff}}}{n_0} \left[|A_x|^4 + \frac{3}{5} |A_y|^4 + \frac{6}{5} |A_x|^2 |A_y|^2 \right] A_x = 0,
\end{aligned}$$

where all phase-dependent nonlinear couplings have been neglected due to waveguide birefringence as justified in Appendix D. The third-order nonlinear terms in equation 6.3 now include vectorial coupling with the orthogonal, transverse field, with Kerr cross-phase modulation coefficient $2\Delta_K$, where $\Delta_K = 1/3$ for linear polarizations for isotropic media under Kleinman symmetry. Optical shock now has vectorial contributions as well, and the Raman cross-coupling coefficients $\Delta_R = \gamma_R = 1/3$. The effective quintic nonlinear index n_4^{eff} consists of three distinct contributions as discussed previously. Because of the additional contributions, the quintic cross-phase modulation coefficients are not in general equal to the intrinsic coefficients due to the symmetry properties of the $\chi^{(5)}$ tensor, but it is presently assumed that this difference is small and the intrinsic values are used.

The previous section studied the propagation of the $\sigma = 0.5$ cubic-quintic eigenmode and showed that stabilized propagation occurred. In this section, the $\sigma = 0.0625$ and the $\sigma = 0.5$ eigenmodes are used in the simulation of the dragging interaction between orthogonally-polarized, initially overlapping solitary waves with slightly different directions of propagation. Two cases are considered in which the $\sigma = 0.0625$ eigenmode is used as both the signal and pump and corresponds to $r = 1.0$, and the $r = 1.7$ case in which the $\sigma = 0.0625$ eigenmode is used as the signal and the $\sigma = 0.5$ eigenmode is used as the pump. In both cases, the gate length is based on the confocal distance of the signal for which $Z_0 = 1.80k_0w_0^2 = 1.65$ cm when $w_0 = 39.6 \mu\text{m}$. The signal spatio-temporal wave has initial intensity spatial FWHM of $81.8 \mu\text{m}$ and temporal FWHM of 33.1 fs, and total energy of 6.37 nJ for a $2 \mu\text{m}$ thick waveguide.

As for the sech()-shaped 1-D spatial solitons, the normalized interaction angle for the spatio-temporal eigenmodes needs to be defined. Generalizing the results of section 5.1.1,

$$\delta k_x = \pi \kappa \left[\frac{c_s}{w_0} + \frac{c_p}{w_p} \right], \quad (6.4)$$

where $\delta k_x = 2\pi \sin\theta/\lambda$ is the transverse wavenumber of the signal, θ is the relative propagation angle between the signal and pump, λ is the material wavelength, $c_s = 0.119$ and $w_0 = 39.6 \mu\text{m}$ for the $\sigma = 0.0625$ eigenmode, and $c_p = 0.301$ and $w_p = w_0$ for the $\sigma = 0.5$ eigenmode. The constants are defined such that $\Delta k_x = 2\pi c_s/w_0$ corresponds to the spatial frequency FWHM (and therefore $\sin\theta = c_s/w_0$ defines the angular FWHM) of the signal, for example, and are determined numerically. The simulations are performed with a 512 by 1024 computational grid which extends from $-427 \mu\text{m}$ to $851 \mu\text{m}$ in the spatial dimension and from -0.520 ps to 1.04 ps in the temporal dimension. In the spatio-temporal frequency domain, the grid extends from -1.26 rad/ μm to 1.25 rad/ μm in the spatial frequency dimension and from -2050 rad/ps (-326 THz) to 2060 rad/ps (328 THz) in the temporal frequency dimension.

The first group of logic gates considered is based on the $r = 1$ dragging interaction with gate lengths up to $10 Z_0$. Before the logic gates themselves can be studied, the stability of the $\sigma = 0.0625$ eigenmode needs to be determined. Figure 6.8 shows the result of propagation of the $\sigma = 0.0625$ pump over $10 Z_0$. For propagation distances of $5 Z_0$ and $10 Z_0$, the spatial broadening factors are 1.34 and 1.89 , respectively. The plot at the bottom of the figure shows the spatial FWHM as a function of propagation distance, where the spatial broadening is due the combination of intra-pulse Raman scattering (leading to temporal broadening) and spatio-temporal coupling, as described in the previous section.

Because of the spatial broadening, the output aperture size must be based on the spatial width of the pump at the output. The aperture size as a function of distance is set to 2.4 times the spatial FWHM; therefore, for a gate length of $5 Z_0$, the aperture size is $264 \mu\text{m}$, and $372 \mu\text{m}$ at gate length $10 Z_0$. This choice of aperture size allows approximately 95% of the pump energy to pass. The throughput of the gate for the undeviated pump is also a function of the Raman downshift, which reduces the energy by converting high frequency photons into low frequency photons. In the case of the $\sigma = 0.0625$ pump, the Raman downshift of 1.90 THz over $10 Z_0$ is relatively small and the energy of the spatio-temporal wave drops from 6.37 nJ to 6.30 nJ.

The variation of threshold contrast with normalized angle and gate length is shown in Figure 6.9. There is a small region in the operating space within which the threshold contrast exceeds the minimum desired value of 10 . A value of the normalized interaction angle $\kappa = 1.2$ is nearly optimal for all gate lengths, and leads to a minimum gate length of about $6.2 Z_0$. Based on

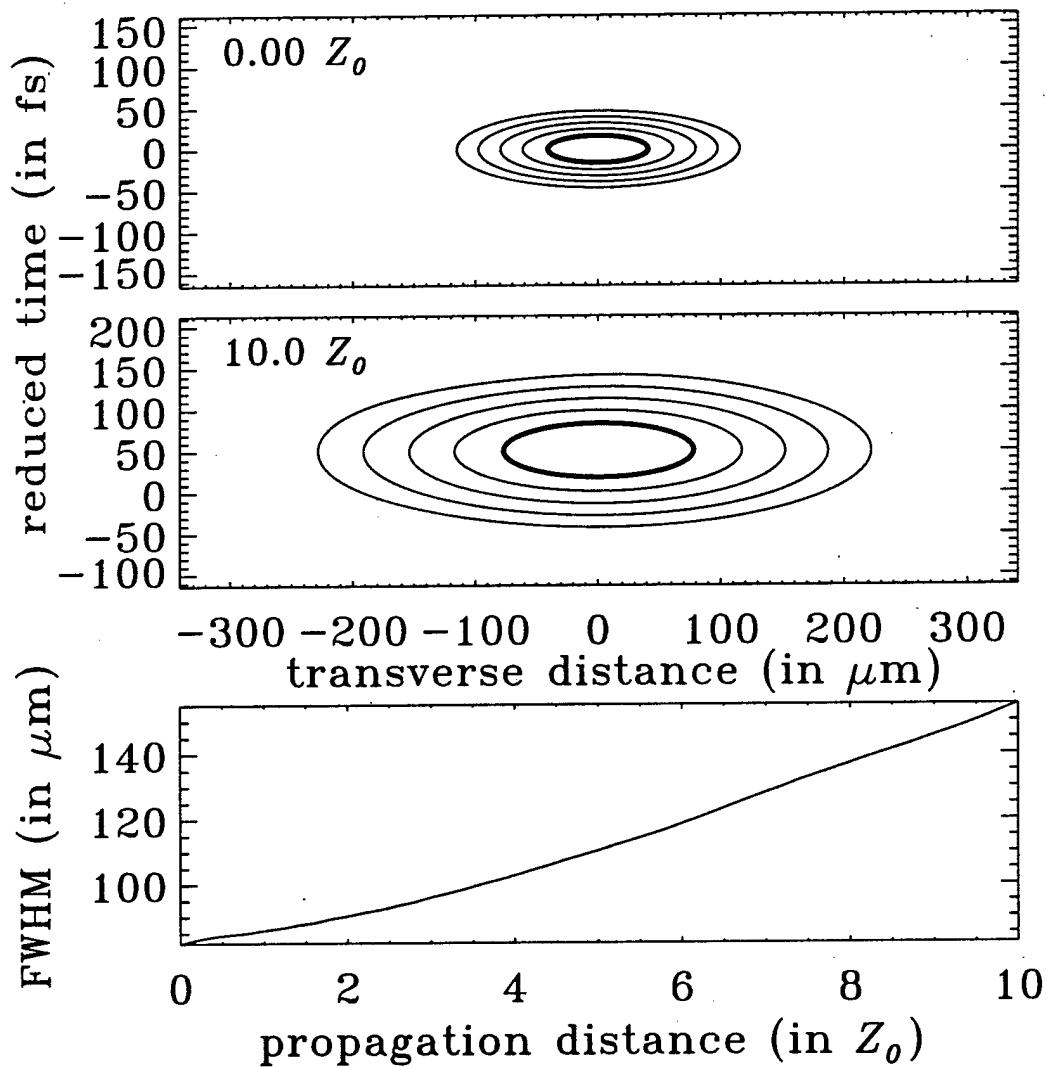


Figure 6.8: Propagation of the $\sigma = 0.0625$ eigenmode over $10 Z_0$. The contour plots show the spatio-temporal profile at the beginning and end of propagation, while the bottom plot shows the evolution of the spatial FWHM with distance. The spatio-temporal wave broadens spatially from $81.8 \mu\text{m}$ at the input to $110 \mu\text{m}$ at $5 Z_0$ and $155 \mu\text{m}$ at $10 Z_0$, and is delayed by 51.1 fs at the output.

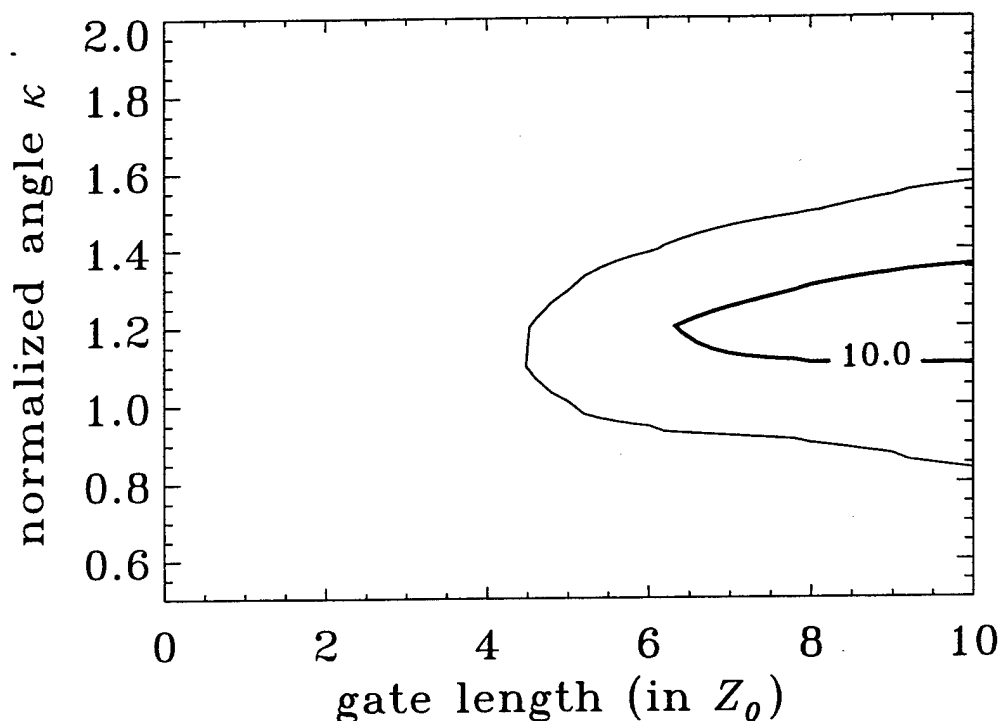


Figure 6.9: Threshold contrast contours versus normalized interaction angle κ and gate length for the $r = 1$ spatio-temporal dragging gate. The minimum gate length satisfying the criterion of threshold contrast ≥ 10 is about $6.2 Z_0$, and the contours are spaced by 5 dB, with the decade contours labeled. The optimal value of the normalized angle is $\kappa \approx 1.2$.

these results, it is expected that the $r = 1$ spatio-temporal dragging gate can provide large-signal gains of $G \geq 1$ for gate lengths greater than $d = 6.2$.

Figure 6.10 shows the spatial dragging interaction between the signal and pump for a normalized interaction angle $\kappa = 1.2$ and gate length of $10 Z_0$. In the first frame, the signal and pump overlap in space and time, with the signal propagating to the right. The second frame (at $5 Z_0$) shows that the signal and pump have formed a bound orbiting pair (indicated by the relative displacement of the centers), with some dispersive radiation escaping the bound potential wells, so that the waves are no longer strictly exponentially localized. The most important point to note from the figure is that sufficient transverse momentum is transferred to the pump during the interaction that very large deflection occurs, such that the -3 dB (at $5 Z_0$) and -15 dB (at $10 Z_0$) contours of the pump lie outside the boundaries of the respective apertures. The threshold contrast metric is 4.39 at $5 Z_0$, and 14.5 at $10 Z_0$.

In addition to the mutual attraction which is responsible for dragging, vectorial nonlinear coupling gives rise to other phenomena as well. Due to cross-focusing, broadening of the pump is reduced considerably when trapped with the signal when compared to propagation alone, with broadening factors of 1.12 and 1.22 at $5 Z_0$ and $10 Z_0$, respectively. Because of cross-shock and cross-Raman effects, the temporal delay and spectral downshift increase. All of these effects act nearly symmetrically on the two waves as shown in the figure because the $r = 1$ pump and signal are identical, except for the fact that the signal is initially propagating at an angle to the propagation axis such that slight off-axis effects such as space-time dispersion manifest themselves.

Figure 6.11 shows the input-output relation for the $r = 1$ spatio-temporal dragging gate of length $10 Z_0$, using $\kappa = 1.2$. Each data point is obtained by calculating the energy of the pump that exits the output aperture as a function of input signal energy. The amplitude of the $\sigma = 0.0625$ signal spatio-temporal wave is multiplied by a constant ranging from 0 to 1.0 (as given by the horizontal axis), in intervals of 0.05. Adjacent calculated data points are averaged to halve the sampling interval in order to increase the placement accuracy of the valid operating regions and the operating points. The most important feature illustrated in the figure is the region of small-signal gain (of about 1.9 at the threshold input level of 0.57) surrounded by saturated levels. These properties give rise to a large-signal gain of 1.05, which is expected to provide a lower bound on the fanout that this gate can achieve in cascaded operation.

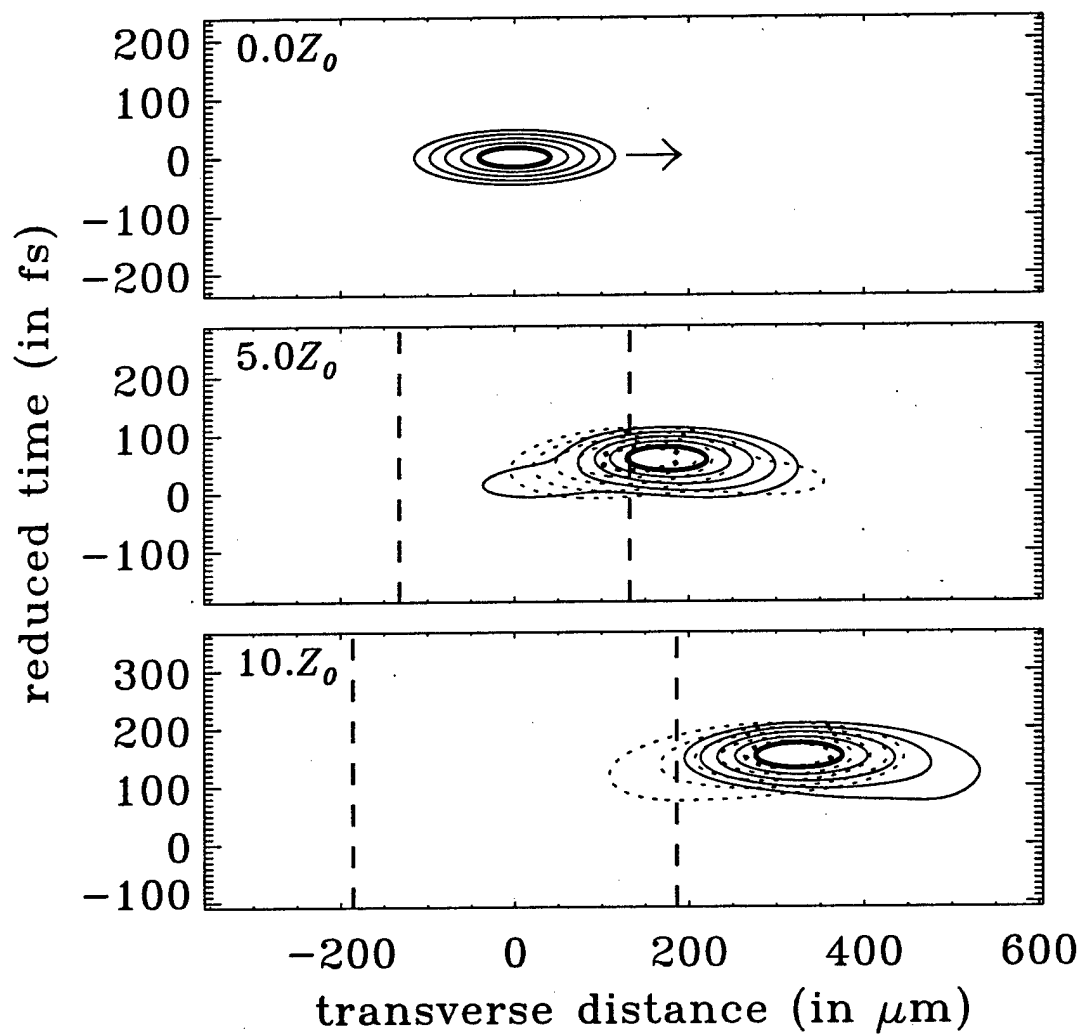


Figure 6.10: The spatial dragging interaction between $\sigma = 0.0625$ fundamental spatio-temporal pump (solid curves) and signal (dotted curves) eigenmodes. The normalized interaction angle $\kappa = 1.2$ and the output aperture width is $264 \mu\text{m}$ at $5 Z_0$ and $372 \mu\text{m}$ $10 Z_0$, indicated by the dashed vertical lines. The threshold contrast is 14.5 for a gate length of $10 Z_0$, and the contours are spaced by 3 dB.

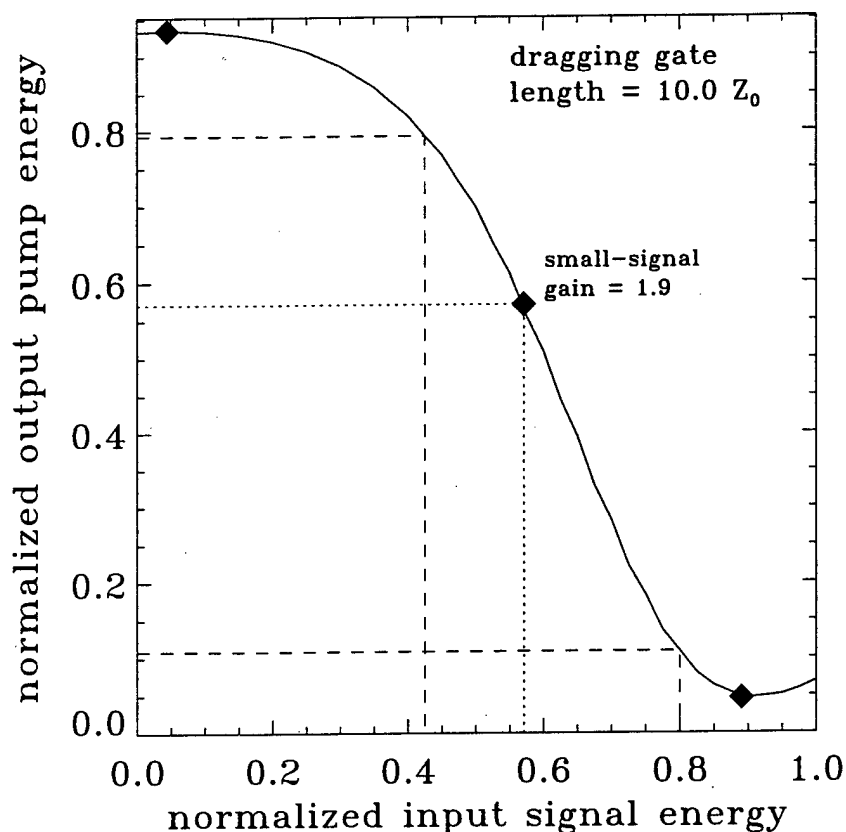


Figure 6.11: Transfer function for the $r = 1$ spatio-temporal dragging gate of length $10 Z_0$, with $\kappa = 1.2$. The threshold contrast is $\tau = 14.5$, and the gate provides large signal gain of $G = 1.05$ with $NM_L = 0.38$ and $NM_H = 0.09$. The high noise margin represents 10% deviation about I_H , and the filled diamonds denote the operating points and threshold.

An additional aspect of the transfer curve of Figure 6.11 worth discussion is that the output level rises above the minimum at the higher input levels. This rise is the result of slight breakup of the pump under symmetric conditions (i.e. unity input level), as illustrated in Figure 6.10. The main effect on gate operation is the reduction of the single-sided noise margin for positive variations about I_H , which is typically assumed to be much larger than the quoted NM_H , which represents the noise margin for negative variations about I_H . For this transfer function, the positive noise margin remains greater than NM_H , because $I_H = 0.89$ and the region of less than unity small-signal gain extends from input levels of 0.8 to more than 1.0. Therefore, the noise performance of this gate is not degraded.

As shown in section 5.4 in the case of 1-D spatial soliton dragging gates, greater large-signal gain is obtained for larger values of r . This trend continues with the spatio-temporal gates as well, as shown by considering the $r = 1.7$ dragging gate. Figure 6.12 shows the propagation of the $\sigma = 0.5$ pump over $10 Z_0$. For gate lengths of $5 Z_0$ and $10 Z_0$, the spatial broadening factors are 1.49 and 2.08, respectively, which are considerably greater than for the $\sigma = 0.0625$ pump, and indicate the stronger influence of Raman scattering on the $\sigma = 0.5$ pump due to increased peak intensity and reduced temporal duration. The plot at the bottom of the figure displays the spatial FWHM as a function of propagation distance, which shows that the most rapid broadening occurs at the beginning of propagation, followed by a plateau, then gradual broadening after about $4.5 Z_0$. Note that this plot corresponds exactly to that of Figure 6.7, given that $Z_0 = 3.21 Z_p$. Between about $2 Z_0$ and $4.5 Z_0$, very little broadening occurs, and the leverage of non-diffracting angular deflection is maximized. Beyond this region, in order for the normalized spatial shift to increase resulting in an increase in threshold contrast, the deflection angle must be greater than the angle of effective broadening.

As before, the aperture size as a function of distance is set to 2.4 times the spatial FWHM; therefore, for a gate length of $5 Z_0$, the aperture size is $145 \mu\text{m}$, and $204 \mu\text{m}$ for gate length of $10 Z_0$. The maximum output energy ratio, or maximum output level, defined as the ratio of the undeflected pump energy exiting the aperture to the fundamental signal eigenmode energy, is a function of aperture size, up to the limit that 100% of the pump energy before the aperture passes through. Because of the

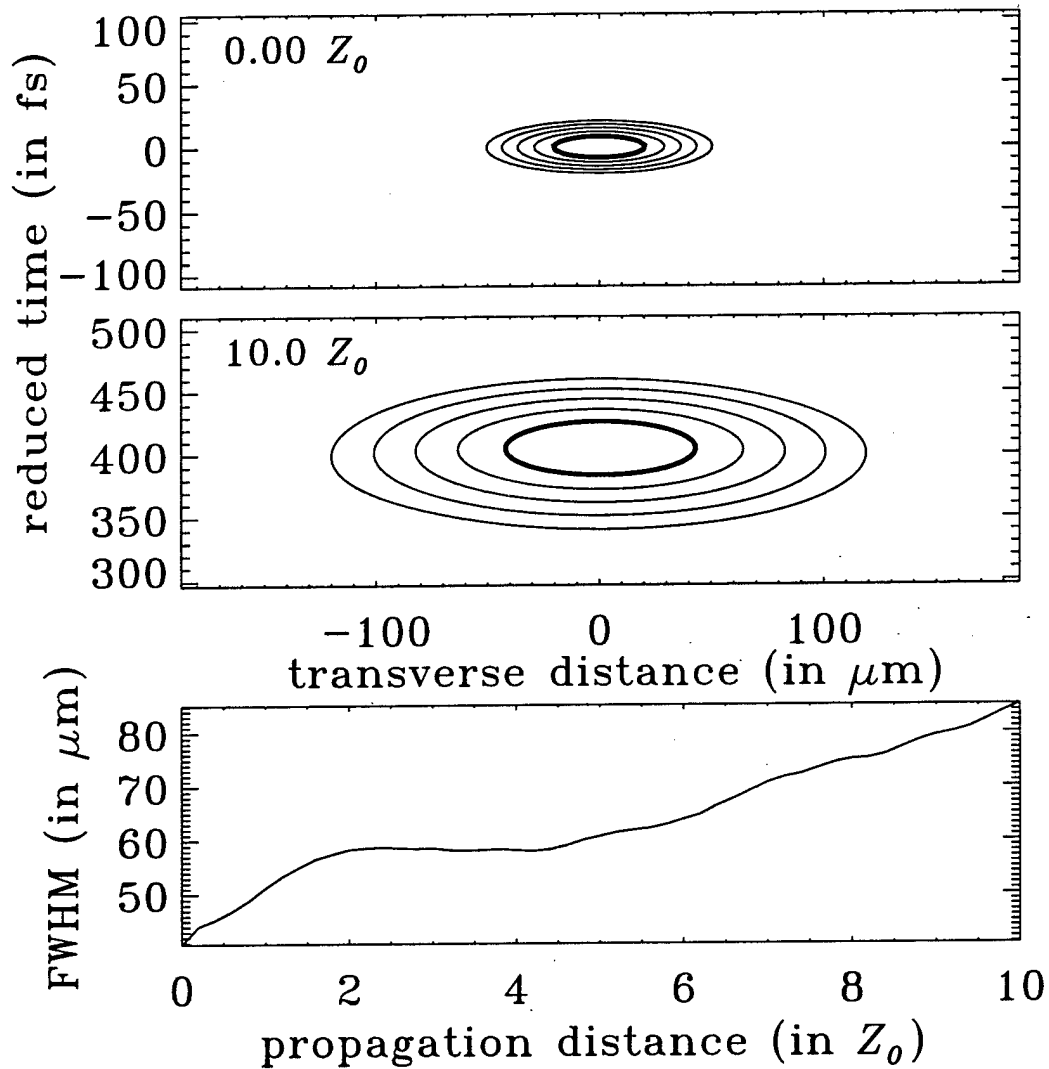


Figure 6.12: Propagation of the $\sigma = 0.5$ pump eigenmode over $10 Z_0$. The contour plots show the spatio-temporal profile at the beginning and end of propagation, while the bottom plot shows the evolution of the spatial FWHM with distance. The spatio-temporal wave broadens spatially from $40.7 \mu\text{m}$ at the input to $60.6 \mu\text{m}$ at $5 Z_0$ and $84.8 \mu\text{m}$ at $10 Z_0$, and is delayed by 391 fs at the output.

Raman downshift, the solitary wave loses energy, but not photon number. In the case of the $\sigma = 0.5$ pump, the Raman downshift of 11.0 THz and 14.4 THz over $5 Z_0$ and $10 Z_0$, respectively, leads to a reduction in energy of the spatio-temporal wave from 10.9 nJ to 10.3 nJ and 10.0 nJ. This optical energy loss may excite optical phonon vibrational modes that can lead to thermal loading of the waveguide even without linear or nonlinear absorption. A smaller aperture size means that a smaller deflection angle for a given gate length is needed to obtain high contrast, but also reduces the maximum output level. The present choice of aperture size allows about 93% of the pump energy to pass. As a result, when using the $\sigma = 0.0625$ signal as the energy reference, the maximum normalized output level is 1.50 at $5 Z_0$ and 1.45 at $10 Z_0$, considerable reductions over the energy ratio of the initial eigenmodes of 1.71.

The variation of threshold contrast with gate length and normalized interaction angle is plotted in Figure 6.13 for the $r = 1.7$ dragging gate. As expected, this interaction produces a much wider range of valid logic gates, with minimum gate length of

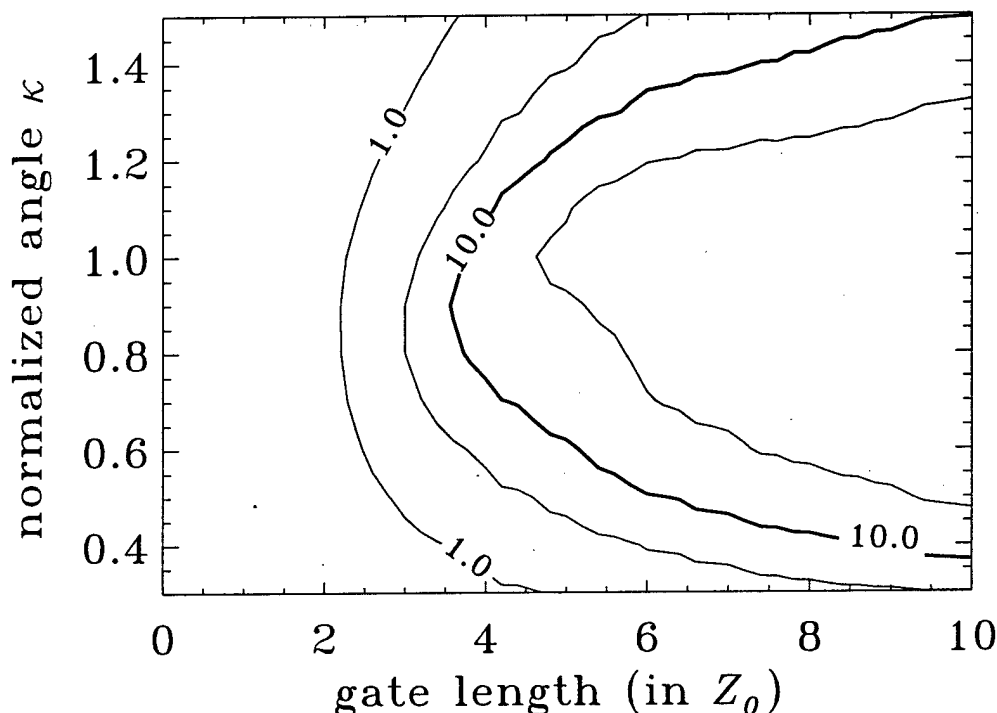


Figure 6.13: Threshold contrast contours versus normalized interaction angle κ and gate length for the $r = 1.7$ spatio-temporal dragging gate. The minimum gate length satisfying the criterion of threshold contrast ≥ 10 is about $3.5 Z_0$, and the contours are spaced by 5 dB, with the decade contours labeled. The optimal value of the normalized angle is $\kappa \approx 1.05$.

about $3.5 Z_0$ for $\kappa \approx 0.9$. Greater large-signal gain than $r = 1.7$ is expected for the gates that lie within the regions of threshold contrast > 10 . Because fanouts of 2 or more are desired, attention will be focused on the $r = 1.7$ dragging gate of lengths $5 Z_0$ or greater, with $\kappa = 1.05$. These gates all lie within the region of threshold contrast > 31.6 , as shown. Complete trapping between the signal and pump occurs for about $\kappa = 0.65$ or less, while dragging occurs for larger angles; therefore, the optimum interaction angle is well within the dragging regime. The figure also shows that high threshold contrast (greater than 10) can be achieved over a large variation of angle, indicating that operation of the spatial dragging gate is relatively insensitive to variations in angle about the optimum value and that critical tolerancing is not required.

Figure 6.14 shows the $r = 1.7$ spatial dragging interaction between the signal and pump for a normalized interaction angle $\kappa = 1.05$ and gate length of $10 Z_0$. In the first frame, the signal and pump overlap in space and time, with the signal propagating to the right. The second frame (at $5 Z_0$) shows that the signal is split into two parts - one that traps with the pump (i.e. the "shadow" [71]), and a dispersive wave that escapes the potential well of the pump. Due to cross-focusing, the trapped portion of the signal also narrows, and is spectrally downshifted in order to equalize group velocity with the pump. Cross-Raman scattering (due to the second Raman term with coefficient Δ_R in equation 6.3) causes additional downshift of the pump as well because of the increased overall intensity in the regions of overlap. In interaction with the signal, the pump delay increases to 144 fs at $5 Z_0$ and 415 fs at $10 Z_0$. This additional timing shift is not detrimental to gate operation though, since the dragged

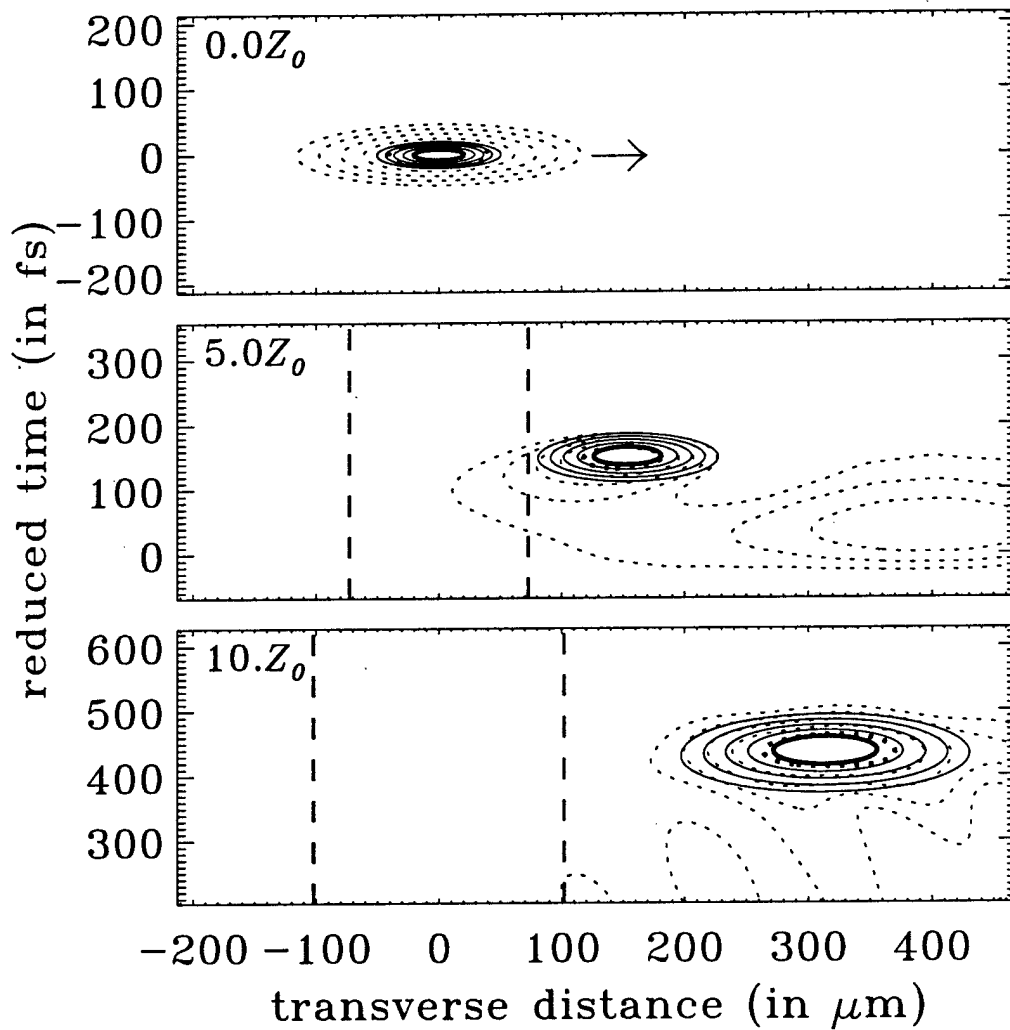


Figure 6.14: The $r = 1.7$ spatial dragging interaction between $\sigma = 0.5$ fundamental spatio-temporal pump (solid curves) and $\sigma = 0.0625$ signal (dotted curves) eigenmodes. The normalized interaction angle $\kappa = 1.05$ and the output aperture width is $146 \mu\text{m}$ at $5 Z_0$ and $204 \mu\text{m}$ $10 Z_0$, indicated by the dashed vertical lines. The threshold contrast is 33.6 and 68.9 for gate lengths of $5 Z_0$ and $10 Z_0$, respectively, and the contours are spaced by 3 dB.

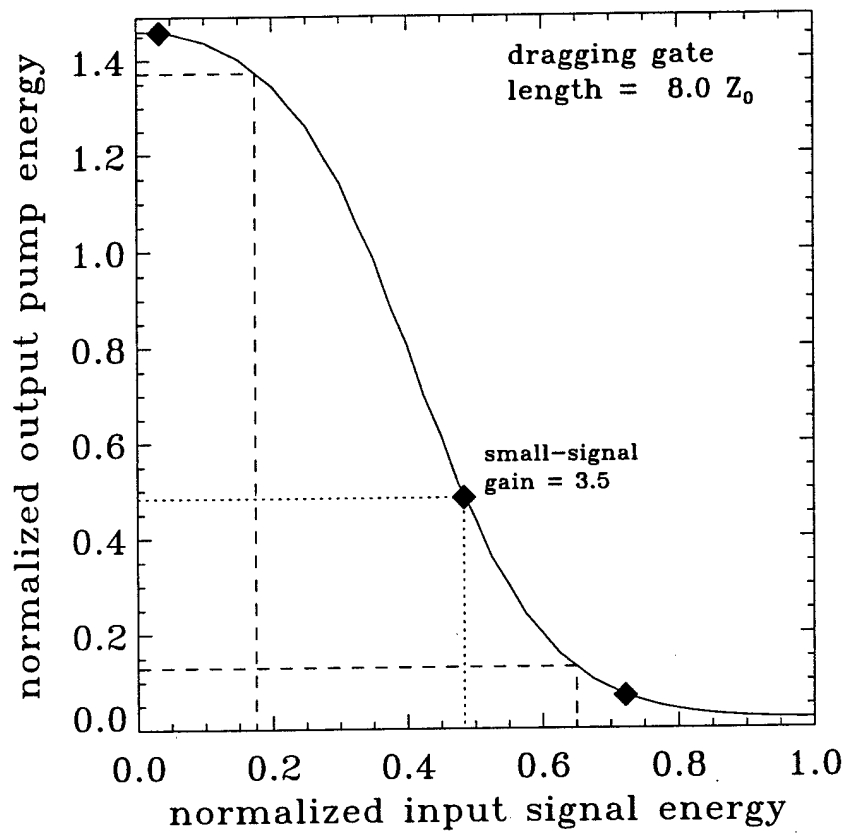


Figure 6.15: Transfer function for the $r = 1.7$ spatio-temporal dragging gate of length $8 Z_0$, with $\kappa = 1.05$. The threshold contrast is $\tau = 49.1$, and the gate provides large signal gain of $G = 2.05$ with $NM_L = 0.14$ and $NM_H = 0.07$. The high noise margin represents 10% deviation about I_H , and the filled diamonds denote the operating points and threshold.

pump is always discarded and never passed on to subsequent stages.

Again, the most important feature from the figure (middle and last frames) is that sufficient transverse momentum is transferred to the pump during the interaction that very large deflection occurs. The threshold contrast metric is 33.6 at $5 Z_0$, and 68.9 at $10 Z_0$, indicating that large-signal gains of at least 1.50 and 1.45, respectively, should be obtainable at each gate length. At $5 Z_0$, the center of the pump solitary wave is well to the right-hand side of the aperture. Beyond the point at which the center of the pump reaches the aperture boundary, because the wave is localized, the threshold metric increases with distance. As a result, the differential change in threshold contrast with gate length can be very large. However, because of spatial broadening of the pump, the aperture size also increases with distance, but large differential change can still be obtained when the deflection angle is greater than the broadening rate, as it is here.

Figure 6.15 shows the transfer function for the $r = 1.7$ spatio-temporal dragging gate of length $8 Z_0$. This gate length is used for the cascaded studies of the next section. Over this distance, the pump alone spatially broadens to $75.1 \mu\text{m}$, and the output aperture width is $180 \mu\text{m}$. As before, adjacent data points from the simulation are averaged to halve the sampling interval. This transfer function also possesses a region of small-signal gain (of 3.5) surrounded by saturated levels. The large-signal gain for this gate is 2.05 with 10% single-sided noise margin about I_H , and is expected to provide a lower bound on the fanout that this gate can achieve in cascaded operation.

Based on the transfer function for each gate length, large-signal gain as a function of gate length is plotted in Figure 6.16. The value for large-signal gain was maximized with the constraint that the single-sided noise margin about I_H is at least 9%. The large-signal gain increases with gate length, but not as quickly as might be expected due to the widening of the aperture with distance. For gate lengths greater than about $7.7 Z_0$, large-signal gains of 2 or greater are obtained. However, over longer gate lengths, the Raman downshift of the pump increases. It is shown in the next section that this downshift is detrimental to cascaded operation, so longer gates will present more difficulties in this capacity. This is why a gate length of $8 Z_0$ is chosen which is close to the minimum for which large-signal gain of 2 is obtained.

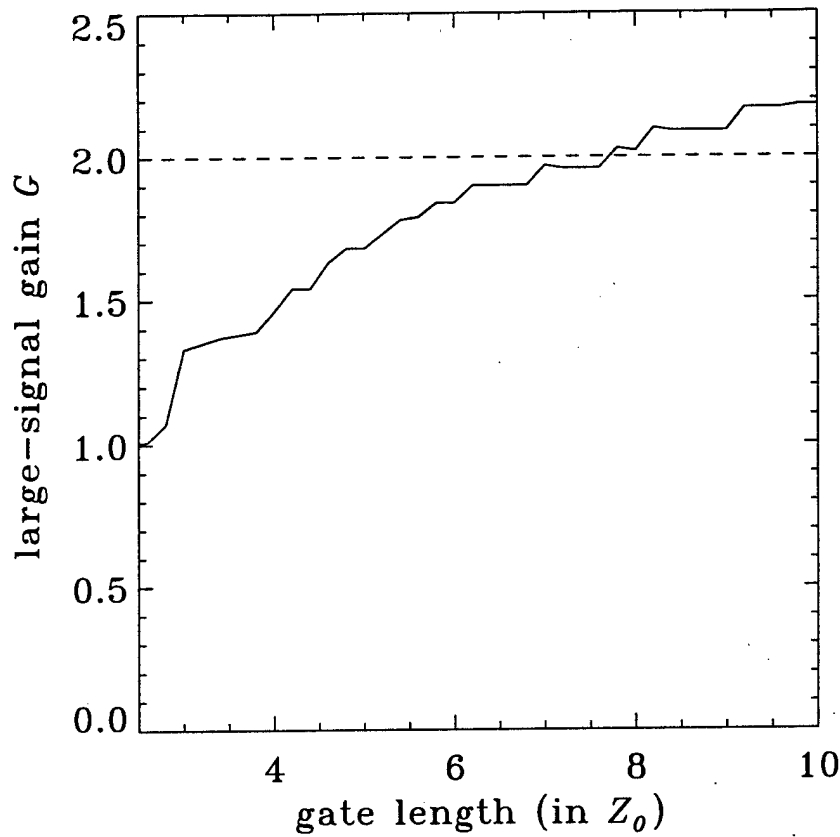


Figure 6.16: Calculated large-signal gain as a function of gate length for the $r = 1.7$ spatio-temporal dragging gate. For each gate length, the stable operating point I_H (which determines the large-signal gain) is chosen to allow for a single-sided noise margin of approximately 9%. Here, the normalized interaction angle $\kappa = 1.05$.

The next section considers the cascaded operation of the $r = 1.7$ spatio-temporal dragging gate of length $8 Z_0$. It is expected from the present results that fanouts of 2 or greater can be obtained with reasonable noise margin.

6.3 Cascaded Spatio-Temporal Dragging Logic Gates

As before, cascaded operation is obtained by taking the pump output of one gate and feeding it into the signal input (ring-oscillator, Figure 5.31) or pump input (multi-input NOR, Figure 5.35) of a subsequent gate. The main issue to deal with in the spatio-temporal case is the strong Raman downshift of the pump. When cascaded to a subsequent stage, there will be a relative downshift, and hence relative group delay, between the signal and pump, leading to temporal walkoff. As originally shown for temporal solitons in fiber [57], two solitons propagating at different group velocities can trap each other and propagate at the weighted mean group velocity. This is the basis for the temporal soliton trapping and dragging logic gates. Temporal dragging/trapping assists in the present situation as well by reducing, or even eliminating, temporal walkoff, but it also reduces the effectiveness of spatial dragging/trapping, leading to reduced spatial shift.

The solution to the temporal walkoff problem is to utilize waveguide birefringence to equalize the initial group velocities of the signal and pump spatio-temporal waves. After propagation over $8 Z_0$, the pump is downshifted by approximately $\Delta\omega = -13$ THz. The change in group delay, evaluated by the expression

$$\Delta k' = \Delta\omega k_0'' + \frac{1}{2}\Delta\omega^2 k_0''' + \frac{1}{6}\Delta\omega^3 k_0''', \quad (6.5)$$

is then $\Delta k' = 2.8$ fs/mm for the fused silica dispersion parameters. Appendix D shows how to achieve this value of differential group delay between the fundamental orthogonal modes using glass waveguide parameters.

An additional effect that needs to be considered is due to the cross-Raman coupling [3, 149]. In the case of the inverter or the first stage of the 2-NOR, the cascaded signal will be downshifted by ~ 13 THz relative to the fresh pump. This downshift nearly corresponds to the peak of the Raman gain. As a result, energy (photons) will be transferred from the pump to the signal through the cross-Raman gain process (described by the third Raman term of equation 6.3, with coefficient γ_R), effectively increasing the threshold contrast, but also increasing the broadening of the pump. The transfer of energy to the signal is also expected to enhance dragging of the pump. Note that these effects are strongest when the group velocities of the pump and signal are equal and when the initial interaction angle is small. Therefore, the optimal interaction angle $\kappa = 1.05$ used in the previous section may no longer be optimal for the cascaded gates.

The following sections discuss the simulation results for two cascaded inverters and for an inverter cascaded into a 2-NOR gate. In each case, fanout and noise margin are calculated, where the effects of Raman scattering are found to be the main limiting factor to the available fanout.

6.3.1 Ring Oscillator Configuration

Due to the long computation time, only a single cascaded inverter driven by the $r = 1.7$ spatio-temporal inverter is considered. As shown in section 5.4, nearly complete logic level restoration is obtained after the first cascaded device, so the second cascaded device need not be considered.

Because of the strong Raman downshift of the pump in the first inverter, the signal in the second inverter will be downshifted (by about 13 THz for gate length of $8 Z_0$) relative to the clean pump. As mentioned previously, this gives rise to two effects. The first is that there will be a difference in group velocities between the interacting waves. The second is that energy will be initially transferred from the (non-downshifted) pump to the (downshifted) signal.

The first effect is countered by equalizing group velocities using the fact that the dispersion properties of the waveguide depend on polarization state, so that the signal propagates in the polarization with the faster group velocity (smaller group delay). The pump still downshifts during propagation, so the net result might be expected to be operation similar to the situation in the beginning of this section in which neither the signal nor pump is initially downshifted. However, the second effect of energy exchange (in addition to asymmetric clipping of the cascaded signal due to the output aperture of the first inverter), alters this expected operation. Like the absorption processes studied in section 5.3, energy transfer from the pump causes additional spatial broadening to occur. This additional broadening means that greater angular deviation is necessary to achieve a resolvable spatial shift (or that reduced spatial shift occurs for the same angular deviation). Unlike the cases studied in section 5.3, the broadening here is non-adiabatic, such that dispersive radiation is generated as the solitary wave reforms, an effect that more than compensates for the fact that less energy is available in the pump to exit the aperture.

The interaction between the pump and downshifted signal over $8 Z_0$ is shown in Figure 6.17. Here, the signal originates from the pump output of a previous gate of length $8 Z_0$ (with no signal input) divided by the fanout factor of $F = 1.5$, resulting in a maximum cascaded input level of 6.22 nJ (or 0.976 in terms of the normalized levels). The interaction angle has been reduced to $\kappa = 0.9$, which was determined through numerical optimization to provide the greatest threshold contrast. During the interaction, the pump energy drops considerably from 10.9 nJ to 7.25 nJ due to energy transfer to the signal, and spatially broadens to $76.5 \mu\text{m}$, slightly larger than the broadening to $75.1 \mu\text{m}$ that would occur for the pump propagating alone over the same distance. Note that, for comparison, during interaction with the $\sigma = 0.0625$ eigenmode in the first inverter stage, the pump narrows at $8 Z_0$ to $72.6 \mu\text{m}$ spatial FWHM due to cross-focusing with the trapped signal, an effect that increases the effective spatial displacement. As a result of the broadening (and dispersive wave generation) due to rapid energy depletion in the second stage, which overcomes the increase in contrast due to pump energy loss, the threshold contrast (relative to the original $\sigma = 0.0625$ signal eigenmode) drops to 38.6, compared with $\tau = 49.1$ for the first-stage interaction with the full 6.37 nJ, non-downshifted $\sigma = 0.0625$ eigenmode.

The transfer curves for this cascaded inverter configuration are shown in Figure 6.18. Only a fanout of 1.5 can be obtained with adequate noise margin. It should be mentioned that fanout is weakly dependent on gate length in this spatio-temporal case because of the widening of the output aperture to follow the spatial broadening of the pump with no signal input. The fanout factor F is less than the large-signal gain of 2.05 (for the $8 Z_0$ gate length) expected from the transfer curve of the first gate, and is mainly the result of the spatial broadening and dispersive wave generation due to the Raman-induced energy exchange mentioned earlier.

As before in the 1-D spatial soliton case, complete logic-level restoration is obtained at the output of the second inverter, because the pump output is simply an amplitude-scaled version of the cascaded signal input, with the same spatial and temporal FWHM, and same downshift. In order to maintain these properties throughout an entire system, the 2-NOR gate must be of total gate length $8 Z_0$ as well, otherwise, the physical parameters of the undeviated pump output (aside from the amplitude) would not be consistent with the standardized signal inputs to the inverter or 2-NOR gates, which are based on the output of an $8 Z_0$ long gate.

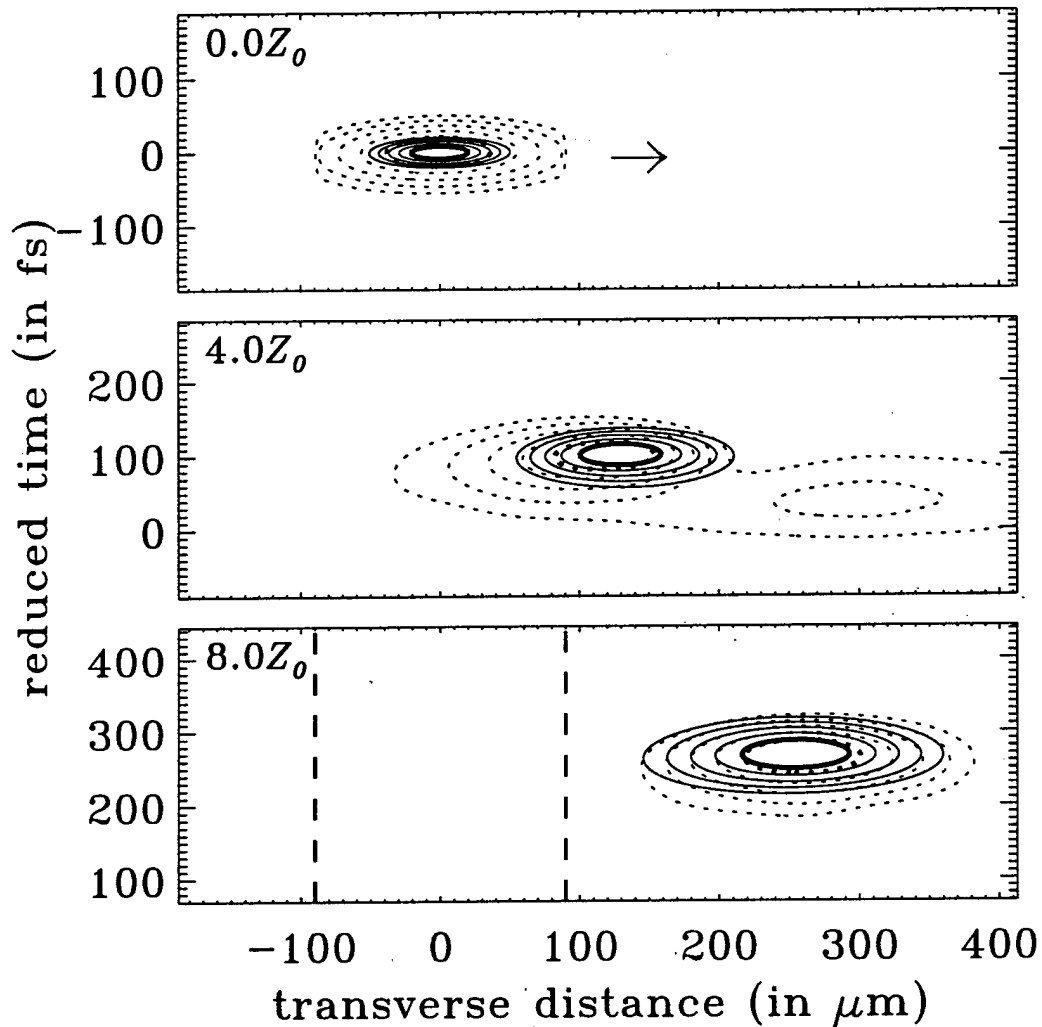


Figure 6.17: The spatial dragging interaction between $\sigma = 0.5$ fundamental spatio-temporal pump (solid curves) and cascaded signal (dotted curves). During the interaction, the pump broadens spatially from $40.7 \mu\text{m}$ to $76.5 \mu\text{m}$, and is reduced in energy from 10.9 nJ to 7.25 nJ , while the signal energy increases from 6.22 nJ to 8.21 nJ (with some unaccounted energy that is absorbed at the computational boundaries). The normalized interaction angle $\kappa = 0.9$ and the output aperture width is $180 \mu\text{m}$ at $8 Z_0$, indicated by the dashed vertical lines. The threshold contrast is 38.6 and the contours are spaced in 3 dB intervals.

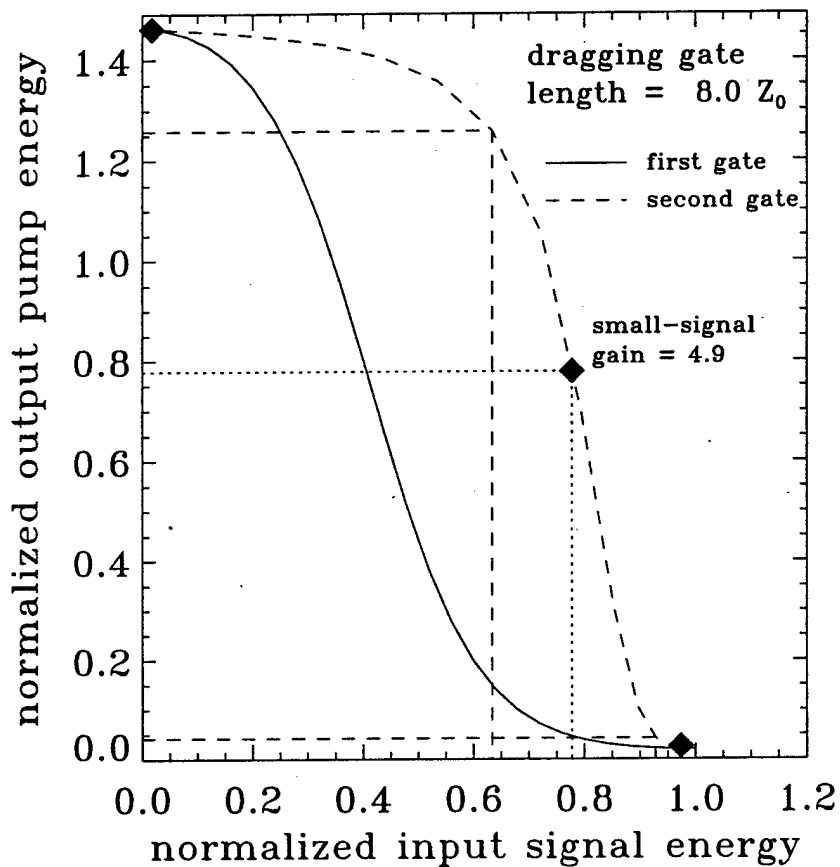


Figure 6.18: Transfer functions for cascaded $r = 1.7$ spatio-temporal dragging gates of length $8 Z_0$, with $\kappa = 1.05$ for the first gate and $\kappa = 0.90$ for the second gate. The threshold contrast for the second inverter is $\tau = 38.6$, and the inverters provide fanout of 1.5 with $NM_L = 0.62$ and $NM_H = 0.05$. The high noise margin represents 4.7% deviation about I_H , and the filled diamonds denote the operating points and threshold.

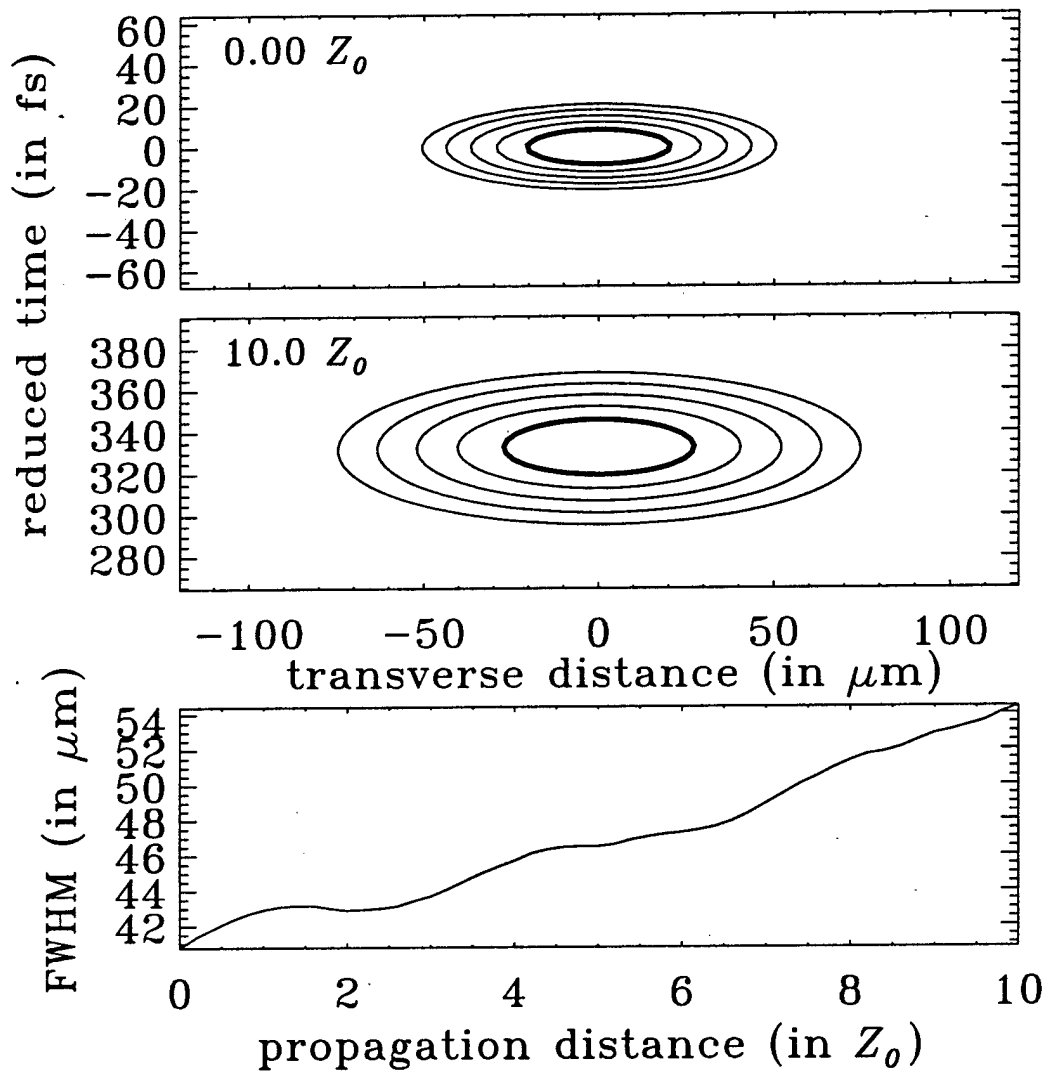


Figure 6.19: Propagation of the $\sigma = 0.5$ pump eigenmode over $10 Z_0$ using reduced Raman nonlinearity. The contour plots show the spatio-temporal profile at the beginning and end of propagation, while the bottom plot shows the evolution of the spatial FWHM with distance. The spatio-temporal wave broadens spatially from $40.7 \mu\text{m}$ at the input to $46.5 \mu\text{m}$ at $5 Z_0$ and $54.4 \mu\text{m}$ at $10 Z_0$, and is delayed by 332 fs at the output.

The detrimental effect of Raman scattering on gate operation can be proven most conclusively by considering the fanout for the cascaded inverters using smaller Raman nonlinearity. The Raman nonlinear coefficient $R_0 = 2.63 \times 10^{-12} \text{ cm}^2/\text{W}\cdot\text{ps}^2$ for fused silica is reduced by a factor of 3.16 to $R_0 = 8.32 \times 10^{-13} \text{ cm}^2/\text{W}\cdot\text{ps}^2$. Using this value, the contribution of the Raman nonlinearity to the total nonlinearity, which becomes 5.5% (with fixed total instantaneous n_2) rather than about 18% for fused silica [150], is more consistent with the ratio for AlGaAs. Reducing the Raman nonlinearity has two effects: the self-induced broadening of the pump is reduced (leading to smaller aperture), and the spectral downshift decreases. The first effect is shown clearly in Figure 6.19 for pump propagation over $10 Z_0$. Now, the spatial broadening of the pump is significantly reduced, with broadening factor 1.34 at $10 Z_0$, rather than 2.08 as in the previous case. For the gate length of $8 Z_0$ used for the calculation of the transfer functions, the spatial broadening is from $40.7 \mu\text{m}$ to $51.4 \mu\text{m}$, a factor of 1.26, leading to an output aperture width of $123 \mu\text{m}$ (versus $180 \mu\text{m}$).

At $8 Z_0$, the spectral downshift of the pump is about 11 THz, leading to a differential group delay of 2.3 fs/mm. The spectral downshift now has about equal contributions from the Raman nonlinearity and the remaining higher-order terms of optical shock and the interplay between positive TOD and SPM. The most important point is that the downshift is now below the peak of the Raman gain, which in addition to the reduction of the Raman nonlinearity, reduces the energy transfer from the pump to the signal for cascaded gates, such that additional broadening of the pump and dispersive wave generation during interaction is greatly reduced.

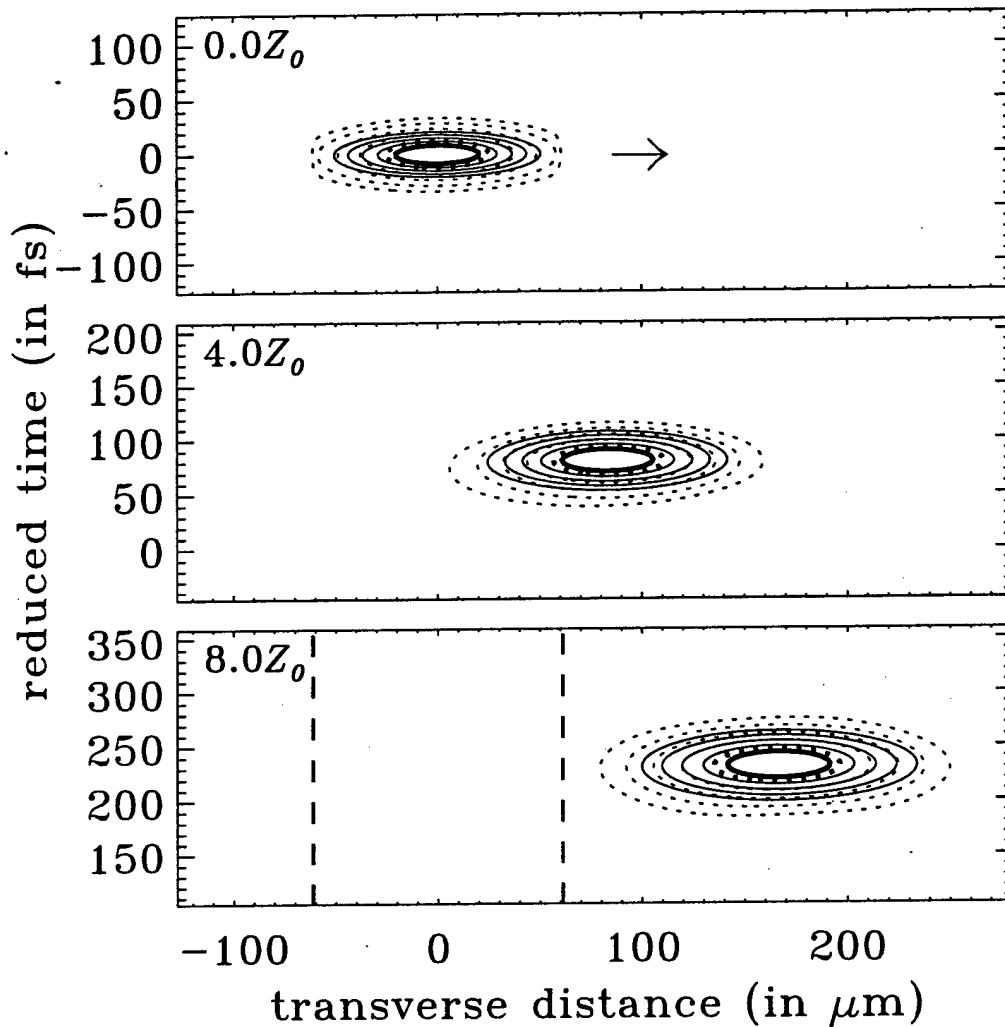


Figure 6.20: The spatial dragging interaction between $\sigma = 0.5$ fundamental spatio-temporal pump (solid curves) and cascaded signal (dotted curves) with reduced Raman nonlinearity. During the interaction, the pump broadens spatially from $40.7 \mu\text{m}$ to $49.4 \mu\text{m}$, and is reduced in energy from 10.9 nJ to 9.32 nJ . The normalized interaction angle $\kappa = 0.95$ and the output aperture width is $123 \mu\text{m}$ at $8 Z_0$, indicated by the dashed vertical lines. The threshold contrast is 236 and the contours are spaced in 3 dB intervals.

The interaction between the pump and downshifted signal, which comes from the full pump output of a previous gate of length $8 Z_0$ divided by the fanout factor of 2.5, is shown in Figure 6.20. During the interaction, the pump energy drops from 10.9 nJ to 9.32 nJ due to energy transfer to the signal, and spatially broadens to $49.4 \mu\text{m}$, slightly less than the broadening to $51.4 \mu\text{m}$ that would occur for the pump propagating alone over the same distance. The reduction in broadening is the result of cross-focusing and increases the effective normalized spatial displacement for a fixed deflection angle. The threshold contrast in this case is $\tau = 236$, which is significantly higher than obtained for the situation of Figure 6.17 with larger proportion of Raman nonlinearity.

In addition to reduced broadening, the reason for the large increase in threshold contrast is that the pump remains exponentially localized down to the -36 dB contour with very little generation of dispersive radiation. For comparison, in the case with larger Raman nonlinearity, the pump remains exponentially localized only within the -24 dB contour. This generation of dispersive radiation is due to the rapid (non-adiabatic) attenuation of the pump such that dispersive energy is shed as the spatio-temporal wave reforms, and is shown in the top plot of Figure 6.21. Most of this dispersive energy propagates on the aperture side of the pump, thereby leading to reduced contrast, because overlap (and hence trapping) with the signal is minimized on that side.

The cascaded transfer functions for the reduced Raman nonlinearity case are plotted in Figure 6.22. Here, a fanout of 2.5

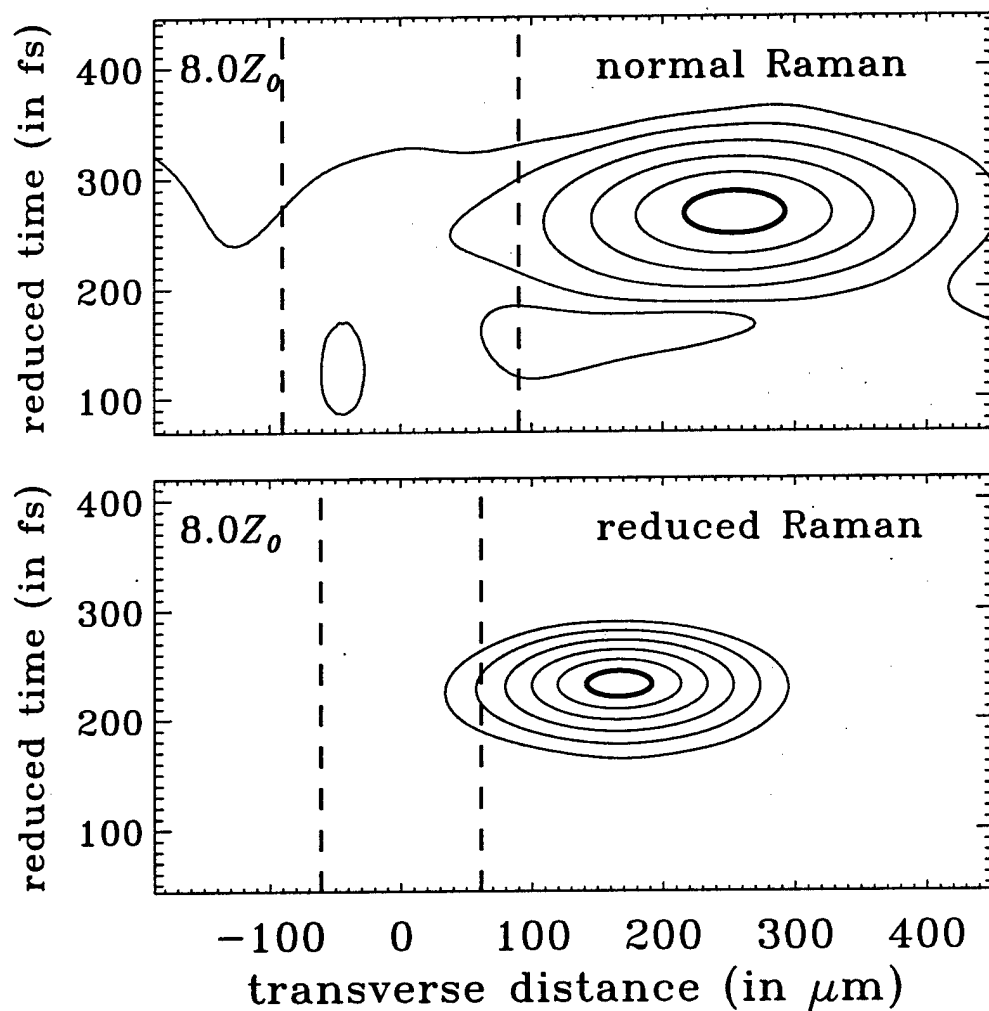


Figure 6.21: Comparison of spatio-temporal pump output intensity profiles for the cascaded $r = 1.7$ inverter gates of length $8 Z_0$ with normal (top) and reduced (bottom) Raman nonlinearity. The first intensity contour (thick linestyle) is at -3 dB relative to each peak, while subsequent contours are spaced by 6 dB, with the last contour being -33 dB relative to the peak. Note that the fanout of the driving gate is different in each case, where $F = 1.5$ (top) and $F = 2.5$ (bottom).

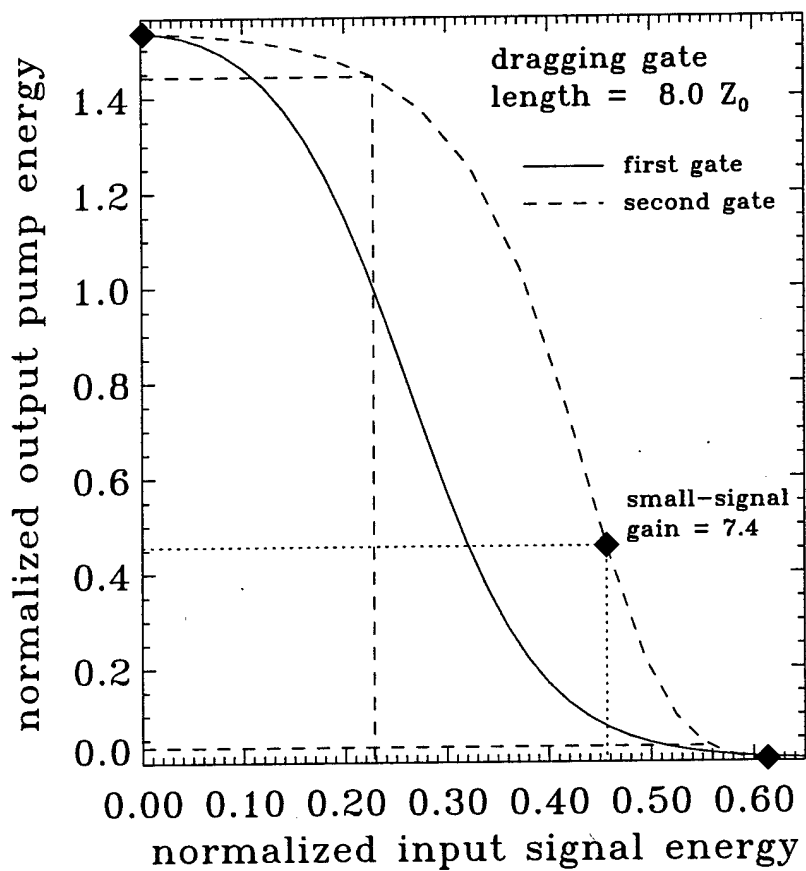


Figure 6.22: Transfer functions for cascaded $r = 1.7$ spatio-temporal dragging gates of length $8 Z_0$ with reduced Raman nonlinearity, with $\kappa = 1.05$ for the first gate and $\kappa = 0.95$ for the second gate. The threshold contrast for the second inverter is $\tau = 236$, and the inverters provide fanout of 2.5 with $NM_L = 0.23$ and $NM_H = 0.05$. The high noise margin represents 8.9% deviation about I_H , and the filled diamonds denote the operating points and threshold.

is obtained with large noise margin representing 9% single-sided deviation about the high operating input level I_H . The most notable differences between these transfer functions and those shown in Figure 6.18 is the reduction in threshold levels for both the first (solid curve) and second inverters (dashed curve). The reason for this reduction is the enhanced localization of the pump during interaction, which keeps dispersive radiation from exiting the aperture, so that the output level increases mainly due to non-resolvable spatial shifts of the pump for low input signal levels.

For the cascaded transfer function, $I_H = 0.61$, meaning that only that fraction of the 6.37 nJ signal energy, or about 4 nJ, is required for switching. In a 100 Gb/s switching system, an energy of 4 nJ per logic operation corresponds to an average power of 400 W, which is well beyond what is reasonable for a practical system, however. As mentioned previously, these results may be more indicative to those obtainable in AlGaAs, for which the ratio of the Raman contribution to the total instantaneous nonlinear refractive index is smaller than for fused silica [151]. In addition, the total nonlinear refractive index of AlGaAs is approximately three-orders of magnitude greater, implying that the switching energy can be reduced by the same factor, to about 4 pJ, with an average power of 0.4 W, which is certainly within the capability of current laser technology.

6.3.2 Two-Input NOR Configuration

As the final device study, the 2-input NOR gate is considered, with inputs into either of two stages driven by the output (divided by the fanout factor) of a single-stage inverter of length $8 Z_0$, as done for the spatial soliton NOR gate of section 5.4.3. Here, only the gate operation with the reduced Raman nonlinearity is studied, which allows for fanout greater than 2. In order to achieve cascability (and signal restoration) with subsequent inverters and NOR gates, the 2-NOR gate must be of length $8 Z_0$. This is to ensure that the spatial width and temporal downshift of each cascaded signal input is the same, so that gates can be stitched together in an arbitrary fashion (within the limits of fanout) without changing gate operation. Note that the results of section 5.4.3 suggest that variations in the input signal width (due to different absorption lengths for the pump between the inverter and 2-NOR gate in the spatial soliton case) are relatively unimportant for consistent gate operation. In the spatio-temporal case, this is also expected to be true, however, the pump downshift does play a big role in cascaded gate operation as discussed in the previous section, and is the main reason why the length of the 2-NOR gate is chosen to equal that of the inverter.

Since the total length of the 2-NOR gate is $8 Z_0$, the sum of the lengths of the two stages must equal this value. The immediate choice is to make the length of each stage $4 Z_0$, using the symmetric geometry of Figure 5.35. Even though the signal is ejected, dragging in the first stage (due to the permanent angle change of the pump) is leveraged over the additional $4 Z_0$ of the second stage such that operation similar to the inverter is expected. However, sufficient dragging by the signal in the second stage is not obtained to achieve fanout of 2 or more from the driving gate. Therefore, it is desirable to increase the length of the second stage, here to $6 Z_0$, for which sufficient fanout is obtained. This asymmetric NOR gate geometry is shown in Figure 5.40. The angle change induced by the signal during the $2 Z_0$ interaction of the first stage still leads to a large spatial shift due to non-diffracting propagation in the second stage, so gate operation from the point of view of signal input 1 is relatively unaffected.

The transfer curves for the spatio-temporal 2-NOR gate are shown in Figure 6.23. Because the dragging length of the second stage is only $6 Z_0$, the fanout of the driving inverter is reduced to $F = 2.25$ in order to achieve adequate noise margin. Here, the single-sided noise margin about I_H for signal input 2 is about 5%, and can be increased by further reducing F of the driving gate. The noise margin obtained for signal input 1 is much greater, about 16%, and is not a limiting factor.

Since the high output state of the 2-NOR gate is virtually identical to the high output state of the inverter, fanouts are expected to be similar as well. Therefore, the 2-NOR gate should be able to drive inverters with $F = 2.5$, and 2-NOR gates with $F = 2.25$. Note that, using the asymmetric implementation of the spatio-temporal 2-NOR gate, the fanout factor of a gate depends on the load that is driven, in contrast to the spatial soliton symmetric 2-NOR gates studied in section 5.4, for which the fanout factor is associated with the particular gate, not the load. The difference lies in the fact that the spatio-temporal inverter and 2-NOR gates are of the same length, so that the high output levels are the same. In addition, the input switching level (which has the minimum value of P_{IH}) of the second input to the 2-NOR gate is higher because the effective gate length is shorter than the inverter. Since the high output level is standardized across all gates, the net result is reduced fanout in order to obtain the higher input switching level for the second stage. In contrast, for the symmetric gate, the output levels of the inverter and 2-NOR gates differed, due to the additional absorption incurred in the second stage of the NOR gate. For these gates, since the 2-NOR consisted of two inverter stages, the input switching levels were the same, so that the fanout is based on the maximum output level of the driving gate, which is lower for the 2-NOR. The asymmetric 2-NOR geometry used here for the spatio-temporal case can be used in spatial case as well, as shown in section 5.4.3, but the reverse is not necessarily true for the spatio-temporal case due to the additional Raman downshift.

As a conclusion to this section, Figure 6.24 illustrates an alternative implementation of the 2-NOR gate, which also applies to the inverter. The main point of this implementation is that the angular deviation of the pump is accomplished in a short

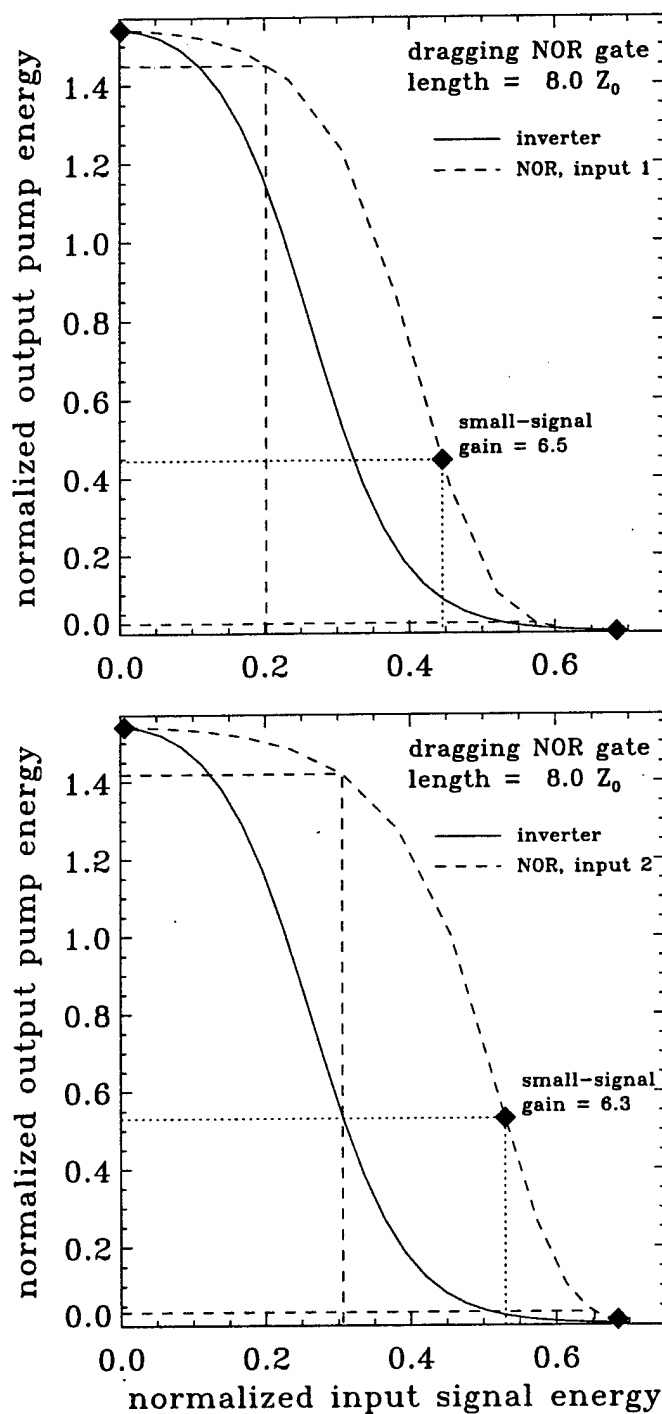


Figure 6.23: Transfer functions (dashed curves) for the cascaded $r = 1.7$ spatio-temporal dragging NOR gate of length $8 Z_0$ with reduced Raman nonlinearity. The 2-NOR gate has two stages of lengths $2 Z_0$ and $6 Z_0$. The first (top) and second (bottom) stage signal inputs are driven by the output of the $8 Z_0$ inverter with fanout factor of 2.25. The threshold contrast of the NOR gate with (worst-case) second stage input is $\tau = 75.7$, and the noise margins $NM_L = 0.30$ and $NM_H = 0.03$. The high noise margin represents 4.7% deviation about I_H .

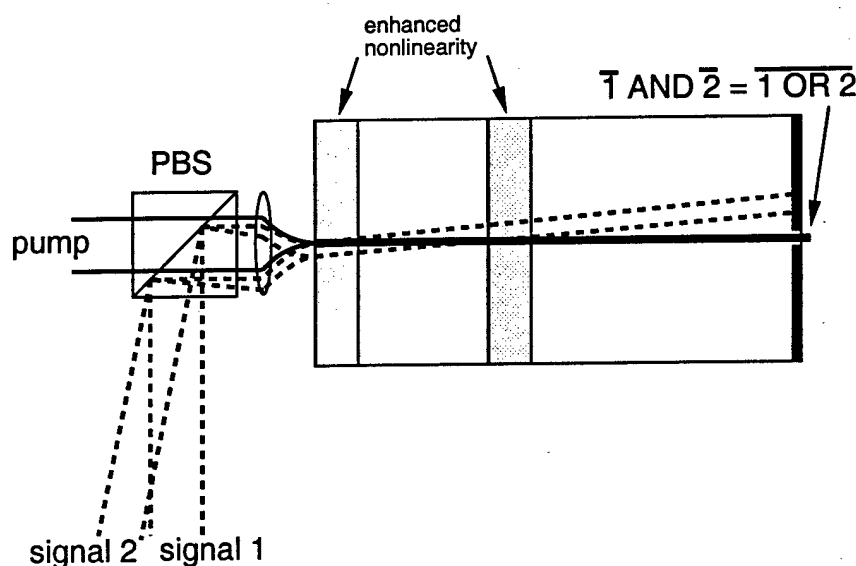


Figure 6.24: Bulk implementation of two-input spatial dragging NOR gate using two asymmetric stages.

section of highly nonlinear material, so that the input signal energy can be reduced. This nonlinearity can be resonant with large absorption because the path length is kept small, about $1 Z_0$, over which the bulk of the angular deflection occurs. Subsequent non-diffracting propagation of the pump then results in a resolvable spatial shift. This geometry is analogous to the generalized time-domain chirp architecture [3] for temporal soliton dragging. An additional aspect of this geometry shown in Figure 6.24 is that a free-space (or linear) region is not needed to separate the two stages of the NOR gate (or any cascaded series of gates), because the asymmetry in the (collision) interaction is provided by the large difference in nonlinear coefficients of the two regions.

6.4 Effects of Material Properties

One of the major conclusions of this chapter is that Raman scattering places the main limitations on optical logic, both from the perspective of the stability of a single spatio-temporal solitary wave, and from the perspective of the cascaded interaction between the pump and downshifted signal waves. It was shown that strong index saturation via $n_4 < 0$ was needed to stabilize the wave against Raman-induced spatio-temporal broadening, which needs to be minimized for the angular deflection geometry to work most effectively. Fused silica may not intrinsically possess the value of n_4^{eff} used for the simulations, however. The use of a long-period longitudinal index grating or waveguide dispersion could be used to adjust the phase velocities of the fundamental and third-harmonic closer to the phase-matched resonance condition in order to generate a large, negative n_4^{casc} . Assuming that this large value of saturation can be obtained, and using the measured Raman response of fused silica, only a fanout of 1.5 was shown in cascaded operation.

Subsequent reduction of the strength of the Raman nonlinearity by about a factor of three allows for increased fanout of 2.5. Further reduction may allow for even greater fanout, but the effects of the other higher-order temporal terms then become prominent. Reduction of the actual material Raman nonlinearity may not be feasible, so it is useful to consider alternate means to effectively achieve this reduction. The higher-order nonlinear temporal effects of Raman scattering and optical shock depend on peak intensity and pulse duration. Therefore, a decrease in either or both of these spatio-temporal wave parameters will reduce these higher-order effects. Decreasing the peak intensity can be accomplished by increasing w_0 , but this also lengthens the logic gate since $Z_0 \propto 1/w_0^2$, and the strength of nonlinear refraction is reduced in proportion, such that no net benefit is obtained. In addition, though, increasing w_0 also increases the pulse duration, which does reduce the effects of Raman and shock relative to nonlinear refraction.

The pulse duration can be increased independently of the intensity (and spatial width) by increasing the magnitude of the group-delay dispersion coefficient, since

$$\tau = w_0 \sqrt{|k_0 k_0''|}, \quad (6.6)$$

and has the benefit of reducing the effects of all higher-order temporal terms without increasing the gate length. However, this has the disadvantage of increasing the energy of the solitary wave in proportion to the pulse duration. Note that the energy is

constant with respect to w_0 (for fixed GDD). These scaling suggest that other material systems with larger values of n_2 should be investigated in order that scaling of the GDD coefficient does not result in spatio-temporal wave energies that exceed practical limitations on average power for high-bandwidth switching systems.

The semiconductor alloy AlGaAs is a natural choice of alternative nonlinear material to fused silica and is considered briefly here. Measurements of the quintic nonlinear index have shown that $n_4^{\text{eff}} \sim -3500 n_2^2$, which is consistent with the value of n_4^{eff} used in the simulations. AlGaAs also has a nonlinear Kerr index 1000 times greater than fused silica, and weaker relative Raman scattering. A recent study [151] suggests that Raman scattering will place a limitation to all-optical switching using short pulses as well, however. The major disadvantage of AlGaAs is that linear dispersion is normal in the transparent spectral window below bandgap. However, this can be overcome with strong waveguide dispersion in the anomalous regime which dominates material dispersion, as demonstrated using ARROW structures [37]. In these structures, AGDD coefficients as large as $k_0'' = -1 \text{ fs}^2/\mu\text{m}$ can be obtained, allowing a wide range of scalings between spatial width and temporal duration in order to minimize the effects of the higher-order temporal terms. In addition, some control over third- and fourth-order dispersion is possible. Another disadvantage is that in the low-loss spectral regime below half-bandgap, nonlinear dispersion is strong due to the proximity to two- and three-photon resonances. This dispersion combined with multi-photon absorption may strongly affect spatio-temporal propagation for pulses with large temporal bandwidth.

Even though there is no clear answer to the materials problem, materials do exist (i.e. fused silica and AlGaAs) in which experiments on spatio-temporal solitary wave propagation and interaction can be performed, and the potential payoff for high bit-rate cascaded logic with fanout justifies such an endeavor. For example, besides small Kerr index, the only major disadvantage to fused silica is the lack of large index saturation. In addition to the methods discussed previously, this can be remedied on two fronts by using very short temporal durations, which not only increases the peak intensity, but begins to freeze out Raman scattering, which cannot build up appreciably over sub-10 fs time scales. Greatly increasing the peak intensity (within the limits of damage threshold of about $1 \times 10^{14} \text{ W/cm}^2$) allows for a smaller value of $n_4^{\text{eff}} < 0$ to achieve the saturation level $\sigma = 0.5$. In addition, with reduced Raman effect, the saturation level σ can be reduced, such that the expected n_4^{casc} of fused silica would provide the necessary stabilization.

Chapter 7

Conclusions

This thesis has developed the necessary framework for the study of general nonlinear, vectorial, spatio-temporal phenomena with large spatial and temporal bandwidths. This framework was applied to the study of novel optical logic devices based on the spatial interaction between 1-D spatial solitons and 2-D spatio-temporal solitary waves. These logic gates were found to have the properties of full level restoration, fanout with large noise margin, and cascability to implement arbitrary logic functionality.

Chapter 1 provided motivation for the study of all-optical switching and logic devices, and covered the basic requirements for these devices as well. The main differences between switching and logic devices is that an optical logic device regenerates degraded data (given sufficient noise margins), has logic-level restoration, provides fanout, and is cascable. Many contemporary switching devices, which do not intrinsically possess these properties, were then discussed: the nonlinear directional coupler, the Kerr gate, the nonlinear optical loop mirror, and the terahertz optical asymmetric demultiplexor. It was then shown that the intrinsic limitations of these devices could be overcome through the use of optical solitons, which propagate without dispersing and/or diffracting. In particular, the temporal and spatial soliton dragging interactions possess the necessary requirements of a three-terminal logic gate, and form the basis for the class of angular deflection logic gate studies of the later chapters, for which a nonlinear phase shift less than π can produce a resolvable change in the output state of the device.

A detailed discussion of optical solitons was given in Chapter 2. Following the historical development of solitary wave and soliton phenomena, 1-D spatial, 1-D temporal, and 2-D spatio-temporal solitons were discussed. In the 1-D cases, soliton solutions were obtained for higher-order equations beyond the traditional (1+1)-D nonlinear Schrödinger equation. For the spatial soliton, this corresponds to the scalar non-paraxial solution, while for the temporal soliton, solutions were shown for equations that contained higher-order dispersive terms one or two orders beyond the slowly-varying envelope approximation (but without Raman scattering). In the temporal soliton section, the effects of third- and fourth-order dispersion, optical shock, and Raman scattering were discussed as well, which have ramifications for the spatio-temporal simulations of Chapter 6. Finally, in the discussion on 2-D spatio-temporal solitary waves, fundamental eigenmode solutions to the (2+1)-D cubic-quintic nonlinear Schrödinger equation were presented, and the stability of these eigenmodes was shown due to the ultrafast saturation effect of a negative quintic nonlinearity. The effects of higher-order perturbations on these eigenmodes were discussed in Chapter 6.

The derivation of the fundamental nonlinear wave equations was presented in Chapter 3. First, the integral representation of the material polarization expansion up to fifth order was transformed into compact differential operator forms suitable for the quasi-monochromatic wave representation. Then, the fully vectorial, nonlinear Helmholtz equation was derived for optical wavepackets centered about a single fundamental frequency. Because of the difficulties this second-order equation presents to analytical and numerical solution, the following section derived via the multiple-scales perturbative technique, a fully vectorial, first-order nonlinear wave equation that consistently includes terms two-orders beyond the slowly-varying amplitude, slowly-varying envelope, and paraxial approximations, in addition to terms that describe the vectorial nonlinear coupling with the weak longitudinally-projected field and nonlinear coupling with a weak third-harmonic wave which produces an effective saturating quintic nonlinearity. This was the fundamental theoretical result of the thesis, and has application not only to the numerical studies of later chapters, but to other areas of study as well. For example, considerable attention has been paid recently to (3+1)-D nonlinear spatio-temporal propagation in the context of pulse splitting in the normal dispersion regime [152, 153]. Simplified NLS-type models [154] have predicted pulse splitting, but it was also recognized that higher-order terms needed to be included in order to follow the subsequent evolution of the split pulses, which have significant angular and spectral bandwidths. More recent work has included scalar space-time non-paraxiality with shock [155], and NLS modified with Raman scattering [153]. These studies are able to qualitatively reproduce the asymmetric splitting behavior observed experimentally [153], but neglect additional physical effects that were derived in this chapter that can also lead to asymmetry.

Chapter 4 discussed linear spatio-temporal diffraction and the split-step numerical method used for the simulations in the remaining chapters. Diffraction in linear homogeneous media can be performed exactly given the correct initial conditions. The diffraction process can be viewed intuitively using the momentum space representation, which is generalized to include spatio-temporal diffraction in dispersive media. The split-step numerical method treats inhomogeneous propagation in two separate steps. The first step is linear homogeneous propagation over a distance short compared to the diffraction length and inhomogeneous length. The second step treats the refractive effect of inhomogeneity due to changes in the linear and/or nonlinear properties of the medium over the same distance. Using the symmetrized representation, the split-step method is expected to be second-order accurate in step-size. However, in the presence of strong nonlinear effects, the accuracy can drop to first-order or less. Even though the scaling can be worse than desirable, comparison of the numerical results to exact analytical and numerical solutions show that sufficient absolute accuracy can be obtained in the simulations.

The heart of the soliton interaction studies for optical logic gates is Chapter 5, which considered logic gates based on spatial soliton interactions. Specifically, the spatial collision and dragging interaction between orthogonally-polarized spatial solitons were found to provide the best performance for logic gates. These interaction were then studied in detail using the threshold contrast metric to find optimal operating parameters, where the dragging interaction generally produced better results. Then, the effects of linear and two-photon absorption on the propagation of a single soliton were studied. Figures-of-merit were derived to evaluate the suitability of a particular nonlinear material for soliton logic applications. The spatial collision and dragging interactions were then evaluated in the presence of absorption with the conclusion that, due to shorter effective interaction distances, the dragging interaction again provided better performance. Finally, using the metrics of small-signal gain, large-signal gain, fanout, and noise margin common in electronics, logic gates based on these interactions were examined. Subsequent cascaded studies show that a sequence of controlled inverters, in which the pump output of one stage divided by the fanout factor serves as the signal input to the next stage, which in the asymptotic limit forms a stable ring oscillator, results in complete logic level restoration and fanouts of two or greater with large noise margin. An additional cascaded geometry, in which the pump output of one stage serves as the pump input to a second stage, implements a two-input NOR gate. This gate was shown to possess complete logic-level restoration as well, with fanouts of two or greater and logic levels compatible with the single-stage inverter. These studies form the second major, systems-level, contribution of this thesis, and are perhaps the first time that such studies have been undertaken for all-optical devices.

The final results chapter, Chapter 6, studied logic gates based on 2-D spatio-temporal solitary waves. Stabilized propagation against the higher-order effects of third- and fourth-order dispersion, space-time focusing, optical shock, and intra-pulse stimulated Raman scattering, was demonstrated due to quintic index saturation. However, it was found that downshifting due to Raman scattering was the most detrimental effect to asymptotically stable propagation. Using this stabilized spatio-temporal wave as the pump, logic gates based on the spatial dragging interaction were studied. The single-stage inverter was shown to provide large-signal gains of two or greater, but cascaded operation proved more difficult due to the strong Raman downshift of the pump which serves as signal inputs to subsequent gates. After equalizing the group velocities of the interacting waves in the cascaded stages, the cross-Raman downshift caused strong energy depletion of the pump, which resulted in spatial broadening and dispersive wave generation, and reduced contrast. Using a weaker proportion of Raman to Kerr nonlinearity (more appropriate for a material such as AlGaAs), fanouts of two or greater were obtained in cascaded operation of inverter and two-input NOR gates with complete logic-level restoration.

The studies of this thesis pave the way for the experimental implementation of low-energy, ultrafast, all-optical logic gates for a variety of applications such as communications switching, routing, and coding, and special-purpose digital computing. The results show that, even though materials issues pose a great challenge for ultrafast systems with low average power, experimental studies using existing materials are warranted and could lead to new capability not possible with current or future electronics.

Bibliography

- [1] J. W. Goodman, F. I. Leonberger, S.-Y. Kung, and R. A. Athale, "Optical interconnections for VLSI systems," *Proceedings of the IEEE*, vol. 72, pp. 850-866, 1984.
- [2] R. McLeod, K. Wagner, and S. Blair, "(3+1)-dimensional optical soliton dragging logic," *Physical Review A*, vol. 52, pp. 3254-3278, October 1995.
- [3] M. N. Islam, *Ultrafast fiber switching devices and systems*. Cambridge University Press, 1992.
- [4] A. E. Willner, "Mining the optical bandwidth for a terabit per second," *IEEE Spectrum*, vol. 34, pp. 32-41, April 1997.
- [5] R. W. Keyes, *The Physics of VLSI Systems*. New York: Addison-Wesley Publishing Co., 1987.
- [6] S. Suzuki, *Photonics in Switching Volume II: Systems*, ch. Time- and wavelength-division switching systems. Academic Press, Inc., 1993.
- [7] H. Takara, S. Kawanishi, M. Saruwatari, and T. Kitoh, "Nearly-penalty-free, fully TDM 100-gbit/s optical transmission by using two tunable mode-locked Er-doped fiber lasers," in *Conference on Optical Fiber Communications, OSA Technical Digest Series*, vol. 4, 1994. paper TuD5.
- [8] L. F. Mollenauer, P. V. Mamyshev, and M. J. Neubelt, "Demonstration of soliton WDM transmission at up to 8×10 gbit/s, error-free over transoceanic distances," in *Optical Fiber Communications*, 1996. Paper PD22.
- [9] M. Nakazawa, K. Suzuki, H. Kubota, and K. Suzuki, "100 gbit/s WDM ($20 \text{ gbit/s} \times 5$ channels) soliton transmission over 1000 km using in-line synchronous modulation and optical filtering," *Electronics Letters*, vol. 33, 1997.
- [10] P. R. Prucnal, "Time-division optical micro-area networks," in *International Conference on Advances in Interconnection and Packaging*, vol. SPIE 1389, 1990.
- [11] K. L. Hall, J. P. Donnelly, S. H. Groves, C. I. Fennelly, R. J. Bailey, and A. Napoleon, "40-Gbit/s all-optical circulating shift register with an inverter," *Optics Letters*, vol. 22, pp. 1479-1481, October 1997.
- [12] P. R. Prucnal, *Photonics in Switching Volume II: Systems*, ch. Photonic fast packet switching. Academic Press, Inc., 1993.
- [13] I. Glesk, K. I. Kang, and P. R. Prucnal, "Ultrafast photonic packet switching with optical control," *Optics Express*, vol. 1, pp. 126-132, September 1997.
- [14] S. D. Smith, I. Janossy, H. A. MacKenzie, J. G. H. Mathew, J. J. E. Reid, M. R. Taghizadeh, F. A. P. Tooley, and A. C. Walker, "Nonlinear optical circuit elements as logic gates for optical computers: the first digital optical circuits," *Optical Engineering*, vol. 24, pp. 569-574, July/August 1985.
- [15] S. D. Smith, "Optical bistability, photonic logic, and optical computation," *Applied Optics*, vol. 25, pp. 1550-1564, May 1986.
- [16] A. Huang, N. Whitaker, H. Avramopoulos, P. French, H. Hough, and I. Chuang, "Sagnac fiber logic gates and their possible applications: a system perspective," *Applied Optics*, vol. 33, pp. 6254-6267, September 1994.
- [17] H. M. Gibbs, S. L. McCall, and T. N. C. Venkatesan, "Differential gain and bistability using a sodium-filled Fabry-Perot interferometer," *Physical Review Letters*, vol. 36, pp. 1135-1138, 1976.
- [18] P. Denyer and D. Renshaw, *VLSI signal processing: a bit-serial approach*. Addison-Wesley Publishing Co., 1985.

- [19] J.-P. Goedgebuer, N. Butterlin, and H. Porte, "Electro-optic systolic processor with a pipeline architecture," *Optics Letters*, vol. 20, pp. 1719-1721, August 1995.
- [20] V. P. Heuring, H. F. Jordan, and J. R. Sauer, "Bit-serial architecture for optical computing," *Applied Optics*, vol. 31, p. 3213, June 1992.
- [21] H. m Lu, "Computational origami: A geometric approach to regular multiprocessing," Master's thesis, Massachusetts Institute of Technology, 1988.
- [22] B. K. Jenkins, P. Chavel, R. Forchheimer, A. A. Sachuck, and T. C. Strand, "Architectural implications of a digital optical processor," *Applied Optics*, vol. 23, pp. 3465-3474, October 1984.
- [23] R. W. Keyes, "Physical limits in digital electronics," *Proceedings of the IEEE*, vol. 63, pp. 740-767, May 1975.
- [24] R. W. Keyes, "What makes a good computer device?," *Science*, vol. 230, pp. 138-144, October 1985.
- [25] D. A. B. Miller, "Device requirements for digital optical processing," *SPIE Critical Reviews: Digital Optical Computing*, vol. CR35, pp. 68-76, 1990.
- [26] L. A. Glasser and D. W. Dobberpuhl, *VLSI Circuits*. Addison-Wesley, 1985.
- [27] P. W. Milonni and J. H. Eberly, *Lasers*. John Wiley & Sons, 1988.
- [28] S. D. Smith, A. C. Walker, F. A. P. Tooley, and B. S. Wherrett, "The demonstration of restoring digital optical logic," *Nature*, vol. 325, pp. 27-31, January 1987.
- [29] P. D. Maker and R. W. Terhune, "Study of optical effects due to an induced polarization third-order in the electric field," *Physical Review A*, vol. 137, p. 801, February 1965.
- [30] G. P. Agrawal, *Nonlinear Fiber Optics*. Academic Press, 1989.
- [31] S. M. Jensen, "The nonlinear coherent coupler," *IEEE Journal of Quantum Electronics*, vol. QE-18, pp. 1580-1583, October 1982.
- [32] G. I. Stegeman and E. M. Wright, "All-optical waveguide switching," *Optical and Quantum Electronics*, vol. 22, March 1990.
- [33] A. Villeneuve, P. Mamyshev, J. U. Kang, G. I. Stegeman, J. S. Aitchison, and C. N. Ironside, "Efficient time-domain demultiplexing with separate signal and control wavelengths in an algaas nonlinear directional coupler," *IEEE Journal of Quantum Electronics*, vol. 31, pp. 2165-2172, December 1995.
- [34] A. M. Weiner, Y. Silberberg, H. Fouckhardt, D. E. Leaird, M. A. Saifi, M. J. Andrejco, and P. W. Smith, "Use of femtosecond square pulses to avoid pulse breakup in all-optical switching," *IEEE Journal of Quantum Electronics*, vol. 25, pp. 2648-2655, December 1989.
- [35] S. Trillo, S. Wabnitz, E. M. Wright, and G. I. Stegeman, "Soliton switching in fiber nonlinear directional couplers," *Optics Letters*, vol. 13, pp. 672-674, August 1988.
- [36] P. Dumais, A. Villeneuve, and J. S. Aitchison, "Bright temporal solitonlike pulses in self-defocusing AlGaAs waveguides near 800 nm," *Optics Letters*, vol. 21, pp. 260-262, February 1996.
- [37] C. J. Hamilton, B. Vögele, J. S. Aitchison, G. T. Kennedy, W. Sibbett, W. Biehlig, U. Peschel, T. Peschel, and F. Lederer, "Bright solitary pulses in AlGaAs waveguides at half the band gap," *Optics Letters*, vol. 21, pp. 1226-1228, August 1996.
- [38] N. J. Doran and D. Wood, "Nonlinear-optical loop mirror," *Optics Letters*, vol. 13, pp. 311-313, January 1988.
- [39] J. S. Aitchison, A. Villeneuve, and G. I. Stegeman, "All-optical switching in two cascaded nonlinear directional couplers," *Optics Letters*, vol. 20, pp. 698-670, April 1995.
- [40] S. Wabnitz, E. M. Wright, C. T. Seaton, and G. I. Stegeman, "Instabilities and all-optical phase-controlled switching in a nonlinear directional coupler," *Applied Physics Letters*, vol. 49, pp. 838-840, October 1986.

- [41] M. A. Duguay and J. W. Hansen, "An ultrafast light gate," *Applied Physics Letters*, vol. 15, pp. 192–194, 1969.
- [42] F. L. Pedrotti and L. S. Pedrotti, *Introduction to Optics*. Prentice Hall, 1987.
- [43] T. Morioka, M. Saruwatari, and A. Takada, "Ultrafast optical multi/demultiplexer utilising optical Kerr effect in polarisation-maintaining single-mode fibers," *Electronics Letters*, vol. 23, pp. 453–454, April 1987.
- [44] O. Dühr, F. Seifert, and V. Petrov, "Ultrafast kerr demultiplexing up to 460 Gbits/s in short optical fibers," *Applied Optics*, vol. 34, no. 24, pp. 5297–5300, 1995.
- [45] M. Jinno and T. Matsumoto, "Nonlinear Sagnac interferometer switch and its applications," *IEEE Journal of Quantum Electronics*, vol. 28, pp. 875–882, April 1992.
- [46] P. A. Andrekson, N. A. Olsson, D. J. Digiovanni, P. A. Morton, T. Tanbun-Ek, R. A. Logan, and K. W. Wecht, "64 Gbit/s all-optical demultiplexing with the nonlinear optical-loop mirror," *IEEE Photonics Technology Letters*, vol. 4, pp. 644–647, June 1992.
- [47] H. Avramopoulos, P. M. W. French, M. C. Gabriel, H. H. Hough, J. N. A. Whitaker, and T. Morse, "Complete switching in a three-terminal Sagnac switch," *IEEE Photonics Technology Letters*, vol. 3, pp. 235–237, March 1991.
- [48] K. Uchiyama, S. Kawanishi, H. Takara, T. Morioka, and M. Saruwatari, "100 Gbit/s to 6.3 Gbit/s demultiplexing experiment using polarization-independent nonlinear optical loop mirror," *Electronics Letters*, vol. 30, pp. 873–875, May 1994.
- [49] X. D. Cao, B. C. Barnett, K. H. Ahn, Y. Liang, G. R. Williams, M. Vaziri, and M. N. Islam, "Experimental cascaded operation of low-birefringence nonlinear-optical loop mirrors," *Optics Letters*, vol. 21, pp. 1211–1213, August 1996.
- [50] J. P. Sokoloff, P. R. Prucnal, I. Glesk, and M. Kane, "A terahertz optical asymmetric demultiplexor (TOAD)," *IEEE Photonics Technology Letters*, vol. 5, pp. 787–790, July 1993.
- [51] M. G. Kane, I. Glesk, J. P. Sokoloff, and P. R. Prucnal, "Asymmetric optical loop mirror: analysis of an all-optical switch," *Applied Optics*, vol. 33, pp. 6833–6842, October 1994.
- [52] I. Glesk, J. P. Sokoloff, and P. R. Prucnal, "Demonstration of all-optical demultiplexing of TDM data at 250 Gbit/s," *Electronics Letters*, vol. 30, pp. 339–340, February 1994.
- [53] K. I. Kang, T. G. Chang, I. Glesk, P. R. Prucnal, and R. K. Boncek, "Demonstration of ultrafast, all-optical, low control energy, single wavelength, polarization independent, cascadable, and integrable switch," *Applied Physics Letters*, vol. 67, pp. 605–607, July 1995.
- [54] N. S. Patel, K. L. Hall, and K. A. Rauschenbach, "40-Gbit/s cascadable all-optical logic with an ultrafast nonlinear interferometer," *Optics Letters*, vol. 21, pp. 1466–1468, September 1996.
- [55] R. J. Manning and G. Sherlock, "Recovery of a π phase shift in ~ 12.5 ps in a semiconductor laser amplifier," *Electronics Letters*, vol. 31, pp. 307–308, February 1995.
- [56] C. R. Menyuk, "Stability of solitons in birefringent optical fibers. I. Equal propagation amplitudes," *Optics Letters*, vol. 12, pp. 614–616, August 1987.
- [57] C. R. Menyuk, "Stability of solitons in birefringent optical fibers. II. Arbitrary amplitudes," *Journal of the Optical Society of America B*, vol. 5, pp. 392–402, February 1988.
- [58] S. Blair, K. Wagner, and R. McLeod, "Asymmetric spatial soliton dragging," *Optics Letters*, vol. 19, pp. 1943–1945, December 1994.
- [59] M. N. Islam, "All-optical cascadable NOR gate with gain," *Optics Letters*, vol. 15, pp. 417–419, April 1990.
- [60] G. Khitrova, H. M. Gibbs, Y. Kawamura, H. Iwamura, T. Ikegami, J. E. Sipe, and L. Ming, "Spatial solitons in a self-focusing semiconductor gain medium," *Physical Review Letters*, vol. 70, pp. 920–923, February 1993.
- [61] F. Lederer and W. Biehlig, "Bright solitons and light bullets in semiconductor waveguides," *Electronics Letters*, vol. 30, pp. 1871–1872, October 1994.

- [62] M. N. Islam, "Ultrafast all-optical logic gates based on soliton trapping in fibers," *Optics Letters*, vol. 14, pp. 1257-1259, November 1989.
- [63] S. R. Friberg, "Soliton fusion and steering by the simultaneous launch of two different-color solitons," *Optics Letters*, vol. 16, pp. 1484-1486, October 1991.
- [64] Y. Kodama and A. Hasegawa, "Effects of initial overlap on the propagation of optical solitons at different wavelengths," *Optics Letters*, vol. 16, pp. 208-210, February 1991.
- [65] M. N. Islam, C. D. Poole, and J. P. Gordon, "Soliton trapping in birefringent optical fibers," *Optics Letters*, vol. 14, pp. 1011-1013, September 1989.
- [66] M. N. Islam, C. E. Socolich, and D. A. B. Miller, "Low-energy ultrafast fiber soliton logic gates," *Optics Letters*, vol. 15, pp. 909-911, August 1990.
- [67] C. R. Menyuk, "Nonlinear pulse propagation in birefringent fibers," *IEEE Journal of Quantum Electronics*, vol. QE-23, pp. 174-176, February 1987.
- [68] M. Shalaby and A. Barthelemy, "Experimental spatial soliton trapping and switching," *Optics Letters*, vol. 16, pp. 1472-1474, October 1991.
- [69] K. Wagner and R. McLeod, "Spatial soliton dragging gates and light bullets," in *OSA Topical Meeting on Optical Computing*, March 1993. Palm Spings, CA.
- [70] S. Blair, K. Wagner, and R. McLeod, "Material figures-of-merit for spatial soliton interactions in the presence of absorption," *Journal of the Optical Society of America B*, vol. 13, pp. 2141-2153, October 1996.
- [71] Q. Wang, P. K. A. Wai, C.-J. Chen, and C. R. Menyuk, "Soliton shadows in birefringent optical fibers," *Optics Letters*, vol. 17, pp. 1265-1267, September 1992.
- [72] J. P. Robinson and D. R. Anderson, "Soliton logic," *Optical Computing and Processing*, vol. 2, pp. 57-61, January 1992.
- [73] V. E. Zakharov and A. B. Shabat, "Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media," *Soviet Physics JETP*, vol. 34, pp. 62-69, January 1972.
- [74] V. I. Karpman and V. V. Solov'ev, "A perturbational approach to the two-soliton system," *Physica D*, vol. 3, pp. 487-502, 1981.
- [75] J. P. Gordon, "Interaction forces among solitons in optical fibers," *Optics Letters*, vol. 8, pp. 596-598, November 1983.
- [76] K. J. Blow and N. J. Doran, "Bandwidth limits of nonlinear (soliton) optical communications systems," *Electronics Letters*, vol. 19, pp. 429-430, May 1983.
- [77] D. Anderson and M. Lisak, "Bandwidth limits due to mutual pulse interaction in optical soliton communications systems," *Optics Letters*, vol. 11, pp. 174-176, 1986.
- [78] L. F. Mollenauer, S. G. Evangelides, and J. P. Gordon, "Wavelength division multiplexing with solitons in ultra-long distance transmission using lumped amplifiers," *Journal of Lightwave Technology*, vol. 9, pp. 362-367, March 1991.
- [79] E. Hecht, *Optics*. Addison-Wesley, second ed., 1987.
- [80] M. Fogiel, ed., *Handbook of Mathematical, Scientific, and Engineering Formulas, Tables, Functions, Graphs, Transforms*. Research and Education Association, 1986.
- [81] S. Chakravarty, M. J. Ablowitz, J. R. Sauer, and R. B. Jenkins, "Multisoliton interactions and wavelength-division multiplexing," *Optics Letters*, vol. 20, pp. 136-138, January 1995.
- [82] S. Chi and S. Wen, "Optical soliton near zero-dispersion regime in Raman pump fiber," *Optics Communications*, vol. 69, January 1989.
- [83] P. A. Andrekson, N. A. Olsson, P. C. Becker, J. R. Simpson, T. Tanbun-Ek, R. A. Logan, and K. W. Wecht, "Observation of multiple wavelength soliton collisions in optical systems with fiber amplifiers," *Applied Physics Letters*, vol. 57, pp. 1715-1717, October 1990.

- [84] D. R. Andersen, D. E. Hooten, J. G. A. Swartzlander, and A. E. Kaplan, "Direct measurement of the transverse velocity of dark spatial solitons," *Optics Letters*, vol. 15, pp. 783-785, July 1990.
- [85] S. R. Skinner, G. R. Allan, D. R. Anderson, and A. L. Smirl, "Dark spatial soliton propagation in bulk ZnSe," *IEEE Journal of Quantum Electronics*, vol. 27, pp. 2211-2219, September 1991.
- [86] T.-T. Shi and S. Chi, "Nonlinear photonic switching by using the spatial soliton collision," *Optics Letters*, vol. 15, pp. 1123-1125, October 1990.
- [87] F. M. Mitschke and L. F. Mollenauer, "Experimental observation of the interaction forces between solitons in optical fibers," *Optics Letters*, vol. 12, pp. 355-357, May 1987.
- [88] F. Reynaud and A. Barthelemy, "Optically controlled interaction between two fundamental soliton beams," *Europhysics Letters*, vol. 12, pp. 401-405, July 1990.
- [89] J. S. Aitchison, A. M. Weiner, Y. Silberberg, D. E. Leaird, M. K. Oliver, J. L. Jackel, and P. W. E. Smith, "Experimental observation of spatial soliton interactions," *Optics Letters*, vol. 16, pp. 15-17, January 1991.
- [90] J.-R. Bian and A. K. Chan, "A nonlinear all-optical switch using spatial soliton interactions," *Microwave and Optical Technology Letters*, vol. 4, pp. 575-580, December 1991.
- [91] M. Shalaby, F. Reynaud, and A. Barthelemy, "Experimental observation of spatial soliton interactions with a $\pi/2$ phase difference," *Optics Letters*, vol. 17, pp. 778-780, June 1992.
- [92] P. L. Chu and C. Desem, "Mutual interaction between solitons of unequal amplitudes in optical fibre," *Electronics Letters*, vol. 21, pp. 1133-1134, November 1985.
- [93] C. Desem and P. L. Chu, "Reducing soliton interaction in single-mode optical fibres," *IEE Proceedings, Part J*, vol. 134, pp. 145-151, June 1987.
- [94] P. A. Andrekson, N. A. Olsson, J. R. Simpson, T. Tanbun-Ek, R. A. Logan, P. C. Becker, and K. W. Wecht, "Soliton collision interaction force dependence on wavelength separation in fibre amplifier based systems," *Electronics Letters*, vol. 26, pp. 1499-1501, August 1990.
- [95] T. Aakjer, J. H. Povlsen, and K. Rottwitt, "Effects of initial overlap in a wavelength-division-multiplexed soliton transmission system," *Optics Letters*, vol. 18, pp. 1908-1910, November 1993.
- [96] R. B. Jenkins, J. R. Sauer, S. Chakravarty, and M. J. Ablowitz, "Data-dependent timing jitter in wavelength-division-multiplexing soliton systems," *Optics Letters*, vol. 20, pp. 1964-1966, October 1995.
- [97] L. Lefort and A. Barthelemy, "All-optical demultiplexing of a signal using collision and waveguiding of spatial solitons," *IEEE Photonics Technology Letters*, vol. 9, pp. 1364-1366, October 1997.
- [98] A. Barthelemy, C. Froehly, S. Maneuf, and F. Reynaud, "Experimental observation of beams' self-deflection appearing with two-dimensional spatial soliton propagation in bulk Kerr material," *Optics Letters*, vol. 17, pp. 844-846, June 1992.
- [99] B. Crosignani, A. Cutolo, and P. DiPorto, "Coupled-mode theory of nonlinear propagation in multimode and single-mode fibers: envelope solitons and self-confinement," *Journal of the Optical Society of America*, vol. 72, pp. 1136-1141, September 1982.
- [100] S. V. Manakov, "On the theory of two-dimensional stationary self-focusing of electromagnetic waves," *Soviet Physics - JETP*, vol. 38, pp. 248-253, February 1974.
- [101] B. A. Malomed and S. Wabnitz, "Soliton annihilation and fusion from resonant inelastic collisions in birefringent optical fibers," *Optics Letters*, vol. 16, pp. 1388-1390, September 1991.
- [102] A. Villeneuve, J. U. Kang, J. S. Aitchison, and G. I. Stegeman, "Unity ratio of cross- to self-phase modulation in bulk AlGaAs and AlGaAs/GaAs MQW waveguides at half the band gap," *Applied Physics Letters*, vol. 67, p. 760, August 1995.
- [103] Q. Wang, P. Wai, C.-J. Chen, and C. R. Menyuk, "Numerical modeling of soliton-dragging logic gates," *Journal of the Optical Society of America B*, vol. 10, pp. 2030-2039, November 1993.

- [104] R. McLeod, K. Wagner, and S. Blair, "Variational approach to orthogonally-polarized optical soliton interaction with cubic and quintic nonlinearities," *submitted to Physica Scripta*, 1998.
- [105] Y. Kodama *Physics Letters A*, vol. 123, p. 276, 1987.
- [106] X. D. Cao and D. D. Meyerhofer, "Soliton collisions in optical birefringent fibers," *Journal of the Optical Society of America B*, vol. 11, pp. 380-385, February 1994.
- [107] X. D. Cao and D. D. Meyerhofer, "All-optical switching by means of collisions of spatial vector solitons," *Optics Letters*, vol. 19, pp. 1711-1713, November 1994.
- [108] J. U. Kang, G. I. Stegeman, and J. S. Aitchison, "One-dimensional spatial soliton dragging, trapping, and all-optical switching in AlGaAs waveguides," *Optics Letters*, vol. 21, pp. 189-191, February 1996.
- [109] J. Millman and A. Grabel, *Microelectronics*. McGraw-Hill, second ed., 1987.
- [110] D. A. B. Miller, "Refractive Fabry-Perot bistability with linear absorption: theory of operation and cavity optimization," *IEEE Journal of Quantum Electronics*, vol. QE-17, pp. 306-311, March 1981.
- [111] E. Caglioti, S. Trillo, S. Wabnitz, and G. I. Stegeman, "Limitations to all-optical switching using nonlinear couplers in the presence of linear and nonlinear absorption and saturation," *Journal of the Optical Society of America B*, vol. 5, pp. 472-482, February 1988.
- [112] G. I. Stegeman, E. M. Wright, N. Finlayson, R. Zanoni, and C. T. Seaton, "Third order nonlinear integrated optics," *Journal of Lightwave Technology*, vol. 6, pp. 953-970, June 1988.
- [113] S. T. Ho, C. E. Socolich, M. N. Islam, W. S. Hobson, A. F. J. Levi, and R. E. Slusher, "Nonlinear spectroscopy near half-gap in bulk and quantum well GaAs/AlGaAs waveguides," *Applied Physics Letters*, vol. 59, pp. 2558-2560, November 1991.
- [114] K. W. DeLong, K. B. Rochford, and G. I. Stegeman, "Effect of two-photon absorption on all-optical guided-wave devices," *Applied Physics Letters*, vol. 55, pp. 1823-1825, October 1989.
- [115] V. Mizrahi, K. W. DeLong, G. I. Stegeman, M. A. Saifi, and M. J. Andrejco, "Two-photon absorption as a limitation to all-optical switching," *Optics Letters*, vol. 14, pp. 1140-1142, October 1989.
- [116] K. W. DeLong and G. I. Stegeman, "Two-photon absorption as a limitation to all-optical waveguide switching in semiconductors," *Applied Physics Letters*, vol. 57, p. 2063, November 1990.
- [117] A. Hasegawa and Y. Kodama, "Signal transmission by optical solitons in monomode fiber," *Proceedings of the IEEE*, vol. 69, pp. 1145-1150, September 1981.
- [118] J. S. Aitchison, A. M. Weiner, Y. Silberberg, M. K. Oliver, J. L. Jackel, D. E. Leaird, E. M. Vogel, and P. W. E. Smith, "Observation of spatial optical solitons in a nonlinear glass waveguide," *Optics Letters*, vol. 15, pp. 471-473, May 1990.
- [119] Y. Silberberg, "Solitons and two-photon absorption," *Optics Letters*, vol. 15, pp. 1005-1007, September 1990.
- [120] J. S. Aitchison, Y. Silberberg, A. M. Weiner, D. E. Leaird, M. K. Oliver, J. L. Jackel, E. M. Vogel, and P. W. E. Smith, "Spatial optical solitons in planar glass waveguides," *Journal of the Optical Society of America B*, vol. 8, pp. 1290-1297, June 1991.
- [121] Y. Chen and J. Atai, "Absorption and amplification of dark solitons," *Optics Letters*, vol. 16, pp. 1933-1935, December 1991.
- [122] A. B. Aceves and J. V. Moloney, "Effect of two-photon absorption on bright spatial soliton switches," *Optics Letters*, vol. 17, pp. 1488-1490, November 1992.
- [123] X. Yang, Y. S. Kivshar, and B. Luther-Davies, "Is two-photon absorption a limitation to dark soliton switching?," *Optics Letters*, vol. 19, pp. 344-346, March 1994.
- [124] D. L. Lee, *Electromagnetic Principles of Integrated Optics*. John Wiley & Sons, 1986. Chapter 7.

- [125] B. L. Lawrence, M. Cha, J. U. Kang, W. Toruellas, G. Stegeman, G. Baker, J. Meth, and S. Etemad, "Large purely refractive nonlinear index of single crystal p-toluene sulfonate (PTS) at 1600 nm," *Electronics Letters*, vol. 30, p. 447, March 1994.
- [126] A. Villeneuve, C. C. Yang, P. G. J. Wigley, G. I. Stegeman, J. S. Aitchison, and C. N. Ironside, "Ultrafast all-optical switching in semiconductor nonlinear directional couplers at half the band gap," *Applied Physics Letters*, vol. 61, p. 147, July 1992.
- [127] K. Al-hemyari, A. Villeneuve, J. U. Kang, J. S. Aitchison, C. N. Ironside, and G. I. Stegeman, "Ultrafast all-optical switching in GaAlAs directional couplers at 1.55 μm without multiphoton absorption," *Applied Physics Letters*, vol. 63, pp. 3562–3564, December 1993.
- [128] C. C. Yang, A. Villeneuve, G. I. Stegeman, and J. S. Aitchison, "Effects of three-photon absorption on nonlinear directional coupling," *Optics Letters*, vol. 17, pp. 710–712, May 1992.
- [129] M. D. Dvorak, W. A. Schroeder, D. R. Anderson, A. L. Smirl, and B. S. Wherrett, "Measurement of the anisotropy of two-photon absorption coefficients in zincblende semiconductors," *IEEE Journal of Quantum Electronics*, vol. 30, pp. 256–268, February 1994.
- [130] B. J. VanZeghbroeck, 1998. ECEN 4375 Class Notes, University of Colorado.
- [131] M. J. Ablowitz and H. Segur, *Solitons and the inverse-scattering transform*, vol. 4. SIAM Studies in Applied Mathematics, 1981.
- [132] A. Hasegawa and F. Tappert, "Transmission of stationary nonlinear optical pulses in dispersive dielectric fiber," *Applied Physics Letters*, vol. 23, 1973.
- [133] P. M. Goorjian and Y. Silberberg, "Numerical simulation of light bullets using the full-vector time-dependent nonlinear Maxwell equations," *Journal of the Optical Society of America B*, vol. 14, pp. 3253–3260, November 1997.
- [134] S. Blair and K. Wagner, "(2+1)-D propagation of spatio-temporal solitary waves including higher-order corrections," *accepted Optical and Quantum Electronics*, 1998.
- [135] Y. Silberberg, "Collapse of optical pulses," *Optics Letters*, vol. 15, pp. 1282–1284, November 1990.
- [136] A. B. Blagoeva, S. G. Dinev, A. A. Dreischuh, and A. Naidenov, "Light bullets formation in bulk media," *IEEE Journal of Quantum Electronics*, vol. 27, pp. 2060–2065, August 1991.
- [137] P. W. Smith, "On the physical limits of digital optical switching and logic elements," *The Bell Systems Technical Journal*, vol. 61, pp. 1975–1993, October 1982.
- [138] D. Shenoy and A. Puri, "Compensation for the soliton self-frequency shift and the third-order dispersion using bandwidth-limited optical gain," *Optics Communications*, vol. 113, pp. 401–406, January 1995.
- [139] K. Hayata, H. Higaki, and M. Koshiba, "Stable light-bullet formation in a Kerr medium: A route to multidimensional solitons in the femtosecond regime," *IEICE Transactions on Electronics*, vol. E78-C, pp. 38–42, January 1995.
- [140] D. N. Christodoulides and R. I. Joseph, "Femtosecond solitary waves in optical fibres," *Electronics Letters*, vol. 20, pp. 659–660, August 1984.
- [141] R. H. Stolen, C. Lee, and R. K. Jain, "Development of the stimulated Raman spectrum in single-mode silica fibers," *Journal of the Optical Society of America B*, vol. 1, pp. 652–657, August 1984.
- [142] J. U. Kang, A. Villeneuve, M. Sheik-Bahae, G. I. Stegeman, K. Al-hemyari, J. S. Aitchison, and C. N. Ironside, "Limitation due to three-photon absorption on the useful spectral range for nonlinear optics in AlGaAs below half bandgap," *Applied Physics Letters*, vol. 65, pp. 147–149, July 1994.
- [143] G. I. Stegeman, A. Villeneuve, J. S. Aitchison, and C. N. Ironside, *Nonlinear Integrated Optics and All-Optical Waveguide Switching in Semiconductors*, p. p. 415. NATO ASI Series 3, 1995.
- [144] J. E. Rothenberg, "Space-time focusing: breakdown of the slowly varying envelope approximation in the self-focusing of femtosecond pulses," *Optics Letters*, vol. 17, pp. 1340–1342, 1992.

- [145] J. P. Gordon, "Theory of the soliton self-frequency shift," *Optics Letters*, vol. 11, pp. 662-664, October 1986.
- [146] A. T. Ryan and G. P. Agrawal, "Spatiotemporal coupling in dispersive nonlinear planar waveguides," *Journal of the Optical Society of America B*, vol. 12, pp. 2382-2389, December 1995.
- [147] B. L. Lawrence, M. Cha, W. E. Toruellas, G. I. Stegeman, S. Etemad, G. Baker, and F. Kajzar, "Measurement of the complex nonlinear refractive index of single crystal p-toluene sulfonate at 1064 nm," *Applied Physics Letters*, vol. 64, p. 2273, May 1994.
- [148] P. B. Corkum, C. Rolland, and T. Srinivasan-Rao, "Supercontinuum generation in gases," *Physical Review Letters*, vol. 57, pp. 2268-2271, November 1986.
- [149] C. R. Menyuk, M. N. Islam, and J. P. Gordon, "Raman effect in birefringent optical fibers," *Optics Letters*, vol. 16, pp. 566-568, April 1991.
- [150] R. H. Stolen, J. P. Gordon, W. J. Tomlinson, and H. A. Haus, "Raman response function of silica-core fibers," *Journal of the Optical Society of America B*, vol. 6, pp. 1159-1166, June 1989.
- [151] Y.-H. Kao, M. N. Islam, J. M. Saylor, R. E. Slusher, and W. S. Hobson, "Raman effect in AlGaAs waveguides for subpicosecond pulses," *Journal of Applied Physics*, vol. 78, pp. 2198-2203, August 1995.
- [152] J. K. Ranka, R. W. Schirmer, and A. L. Gaeta, "Observation of pulse splitting in nonlinear dispersive media," *Physical Review Letters*, vol. 77, pp. 3783-3786, October 1996.
- [153] S. A. Diddams, H. K. Eaton, A. A. Zozulya, and T. S. Clement, "Characterizing the nonlinear propagation of femtosecond pulses in bulk media," *submitted to IEEE Journal of Selected Topics in Quantum Electronics*, 1997.
- [154] J. E. Rothenberg, "Pulse splitting during self-focusing in normally dispersive media," *Optics Letters*, vol. 17, pp. 583-585, April 1992.
- [155] G. Fibich and G. C. Papanicolaou, "Self-focusing in the presence of small time dispersion and nonparaxiality," *Optics Letters*, vol. 22, pp. 1379-1381, September 1997.
- [156] M. A. Newhouse, D. L. Weidman, and D. W. Hall, "Enhanced-nonlinearity single-mode lead silicate optical fiber," *Optics Letters*, vol. 15, pp. 1185-1187, November 1990.
- [157] M. Asobe, K. Naganuma, T. Kaino, T. Kanamori, S. Tomaru, and T. Kurihara, "Switching energy limitation in all-optical switching due to group-velocity dispersion in highly nonlinear waveguides," *Applied Physics Letters*, vol. 64, pp. 2922-2924, May 1994.
- [158] M. Thakur, R. C. Frye, and B. I. Greene, "Nonresonant absorption of single-crystal films of polydiacetylene measured by photothermal deflection spectroscopy," *Applied Physics Letters*, vol. 56, pp. 1187-1188, March 1990.
- [159] F. Wijnands, T. Rasmussen, H. J. W. M. Hoekstra, J. Povlsen, A. Bjarklev, and R. M. D. Ridder, "Efficient interface conditions for the semi-vectorial finite-difference beam propagation method," *Optical and Quantum Electronics*, vol. 27, pp. 961-975, 1995.
- [160] J. W. Fleming, "Material dispersion in lightguide glasses," *Electronics Letters*, vol. 14, pp. 326-328, May 1978.
- [161] C. S. Gardner, J. M. Greene, M. J. Kruskal, and R. M. Miura, "Method for solving the Korteweg-deVries equation," *Physical Review Letters*, vol. 19, pp. 1095-1097, November 1967.
- [162] M. J. Ablowitz, M. D. Kruskal, A. C. Newell, and H. Segur, "The inverse scattering transform - Fourier analysis for nonlinear problems," *Studies in Applied Mathematics*, vol. 53, pp. 249-315, 1974.
- [163] M. J. Ablowitz and P. A. Clarkson, *Solitons, Nonlinear Evolution Equations, and the Inverse-Scattering Transform*, vol. London Mathematical Society Lecture Note Series 149. Cambridge University Press, 1991.
- [164] J. S. Russell, "Report on waves," in *Rep. 14th Meet. Brit. Assoc. Adv. Sci.*, pp. 311-390, 1844.
- [165] P. G. Drazin and R. S. Johnson, *Solitons: an introduction*. Cambridge University Press, 1989.

- [166] N. J. Zabusky and M. D. Kruskal, "Interactions of 'solitons' in a collisionless plasma and the recurrence of initial states," *Physical Review Letters*, vol. 15, pp. 240-243, August 1965.
- [167] D. J. Korteweg and G. deVries, "On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves," *Philos. Mag.*, vol. 39, no. 5, pp. 422-423, 1895.
- [168] E. Fermi, J. Pasta, and S. M. Ulam, "Studies in nonlinear problems," *Technical Report, Los Alamos Scientific Laboratory*, vol. LA-1940, 1955.
- [169] R. Y. Chiao, E. Garmire, and C. H. Townes, "Self-trapping of optical beams," *Physical Review Letters*, vol. 13, pp. 479-482, October 1964.
- [170] N. F. Pilipetskii and A. R. Rustamov, "Observation of self-focusing of light in liquids," *JETP Letters*, vol. 2, pp. 55-56, 1965.
- [171] M. Hercher, "Laser-induced damage in transparent media," *Journal of the Optical Society of America*, vol. 54, April 1964.
- [172] C. R. Giuliano and J. H. Marburger, "Observation of moving self-foci in sapphire," *Physical Review Letters*, vol. 27, pp. 905-908, October 1971.
- [173] J. E. Bjorkholm and A. Ashkin, "cw self-focusing and self-trapping of light in sodium vapor," *Physical Review Letters*, vol. 32, pp. 129-132, January 1974.
- [174] P. Lallemand and N. Bloembergen, "Self-focusing of laser beams and stimulated Raman gain in liquids," *Physical Review Letters*, vol. 15, pp. 1010-1012, December 1965.
- [175] P. L. Kelley, "Self-focusing of optical beams," *Physical Review Letters*, vol. 15, pp. 1005-1008, December 1965.
- [176] J. J. Rasmussen and K. Rypdal, "Blow-up in Nonlinear Schrödinger Equations - I: A general review," *Physica Scripta*, vol. 33, pp. 481-497, 1986.
- [177] W. G. Wagner, H. A. Haus, and J. H. Marburger, "Large-scale self-trapping of optical beams in the paraxial ray approximation," *Physical Review*, vol. 175, pp. 256-266, November 1968.
- [178] V. I. Bespalov and V. I. Talanov, "Filamentary structure of light beams in nonlinear liquids," *JETP Letters*, vol. 3, pp. 307-310, June 1966.
- [179] J. H. Marburger and E. Dawes, "Dynamical formation of a small-scale filament," *Physical Review Letters*, vol. 21, no. 8, pp. 556-558, 1968.
- [180] H. A. Haus, "Higher order trapped light beam solutions," *Applied Physics Letters*, vol. 8, pp. 128-129, March 1966.
- [181] T. K. Gustafson, P. L. Kelley, R. Y. Chiao, and R. G. Brewer, "Self-trapping in media with saturation of the nonlinear index," *Applied Physics Letters*, vol. 12, pp. 165-168, March 1968.
- [182] L. F. Mollenauer, R. H. Stolen, and J. P. Gordon, "Experimental observation of picosecond pulse narrowing and solitons in optical fibers," *Physical Review Letters*, vol. 45, pp. 1095-1098, September 1980.
- [183] T. K. Gustafson, J. P. Taran, H. A. Haus, J. R. Lifshitz, and P. L. Kelley, "Self-modulation, self-steepening, and spectral development of light in small-scale trapped filaments," *Physical Review*, vol. 177, pp. 306-313, January 1969.
- [184] D. Andersen and M. Lisak, "Nonlinear asymmetric self-phase modulation and self-steepening of pulses in long optical waveguides," *Physical Review A*, vol. 27, pp. 1393-1398, March 1983.
- [185] Y. R. Shen and N. Bloembergen, "Theory of stimulated Brillouin and Raman scattering," *Physical Review*, vol. 137, pp. 1787-1805, March 1965.
- [186] K. J. Blow and D. Wood, "Theoretical description of transient stimulated Raman scattering in optical fibers," *IEEE Journal of Quantum Electronics*, vol. 25, pp. 2665-2673, December 1989.
- [187] F. M. Mitschke and L. F. Mollenauer, "Discovery of the soliton self-frequency shift," *Optics Letters*, vol. 11, pp. 659-661, October 1986.

- [188] A. Barthelemy, S. Maneuf, and C. Froehly, "Propagation soliton et auto-confinement de faisceaux laser par nonlinearite optique de Kerr," *Optics Communications*, vol. 55, no. 3, pp. 201-206, 1985.
- [189] S. Maneuf, R. Desailly, and C. Froehly, "Stable self-trapping of laser beams: Observation in a nonlinear planar waveguide," *Optics Communications*, vol. 65, no. 3, pp. 193-198, 1988.
- [190] J. S. Aitchison, K. Al-Hemyari, C. N. Ironside, R. S. Grant, and W. Sibbett, "Observation of spatial solitons in AlGaAs waveguides," *Electronics Letters*, vol. 28, pp. 1879-1880, September 1992.
- [191] A. Barthelemy, B. Colombeau, C. Froehly, and M. Vampouille, "Stable soliton and nonsoliton three-dimensional propagation of intense light by multiple beams interference in Kerr materials," in *Technical Digest of the OSA 1990 Annual Meeting*, 1990.
- [192] T. K. Gustafson, J. P. E. Taran, P. L. Kelley, and R. Y. Chiao, "Self-modulation of picosecond pulses in electro-optic crystals," *Optics Communications*, vol. 2, pp. 17-21, May 1970.
- [193] J.-M. R. Thomas and J.-P. E. Taran, "Pulse distortions in mismatched second harmonic generation," *Optics Communications*, vol. 4, pp. 329-334, January 1972.
- [194] A. C. Newell and J. V. Moloney, *Nonlinear Optics*. Addison-Wesley, 1992.
- [195] M. J. Ablowitz, G. Biondini, and S. Blair, "Multi-dimensional pulse propagation in non-resonant $\chi^{(2)}$ materials," *Physics Letters A*, vol. 236, pp. 520-524, December 1997.
- [196] R. DeSalvo, D. J. Hagan, M. Sheik-Bahae, G. Stegeman, and E. W. V. Stryland, "Self-focusing and self-defocusing by cascaded second-order effects in KTP," *Optics Letters*, vol. 17, pp. 28-30, January 1992.
- [197] L. Torner, C. R. Menyuk, and G. I. Stegeman, "Excitation of solitons with cascaded $\chi^{(2)}$ nonlinearities," *Optics Letters*, vol. 19, pp. 1615-1617, October 1994.
- [198] A. V. Buryak and Y. S. Kivshar, "Spatial optical solitons governed by quadratic nonlinearity," *Optics Letters*, vol. 19, pp. 1612-1614, October 1994.
- [199] R. Schiek, Y. Baek, and G. I. Stegeman, "One-dimensional spatial solitary waves due to cascaded second order nonlinearities in planar waveguides," *Physical Review E*, vol. 53, January 1996.
- [200] R. G. Wenzel, J. L. Carlsten, and K. J. Drühl, "Soliton experiments in stimulated Raman scattering," *Journal of Statistical Physics*, vol. 39, no. 5/6, pp. 621-632, 1985.
- [201] A. S. Gouveia-Neto, A. S. L. Gomes, and J. R. Taylor, "Femtosecond soliton Raman generation," *IEEE Journal of Quantum Electronics*, vol. 24, pp. 332-340, February 1988.
- [202] J. G. A. Swartzlander and C. T. Law, "Optical vortex solitons observed in Kerr nonlinear media," *Physical Review Letters*, vol. 69, pp. 2503-2506, October 1992.
- [203] A. Braun, G. Korn, X. Liu, D. Du, J. Squier, and G. Mourou, "Self-channeling of high-peak-power femtosecond laser pulses in air," *Optics Letters*, vol. 20, pp. 73-75, January 1995.
- [204] M. Segev, B. Crosignani, A. Yariv, and B. Fischer, "Spatial solitons in photorefractive media," *Physical Review Letters*, vol. 68, pp. 923-926, February 1992.
- [205] N. Belanger, A. Villeneuve, and J. S. Aitchison, "Solitonlike pulses in self-defocusing AlGaAs waveguides," *Journal of the Optical Society of America B*, vol. 14, pp. 3003-3012, November 1997.
- [206] D. N. Christodoulides and R. I. Joseph, "Femtosecond solitary waves in optical fibres - beyond the slowly varying envelope approximation," *Applied Physics Letters*, vol. 47, pp. 76-78, July 1985.
- [207] P. L. Francois, J. F. Bayon, and F. Alard, "Design of monomode quadruple-clad fibres," *Electronics Letters*, vol. 20, pp. 688-689, August 1984.
- [208] Y. Kodama, "On integrable systems with higher order corrections," *Physics Letters*, vol. 107A, pp. 245-249, February 1985.

- [209] A. A. Kolokolov, "Stability of the dominant mode of the nonlinear wave equation in cubic media," *Journal of Applied Mechanics and Technical Physics*, vol. 3, p. 426, 1973.
- [210] D. E. Edmundson and R. H. Enns, "Fully three-dimensional collisions of bistable light bullets," *Optics Letters*, vol. 18, pp. 1609-1611, October 1993.
- [211] S. Saltiel, S. Tanev, and A. D. Boardman, "High-order nonlinear phase shift caused by cascaded third-order processes," *Optics Letters*, vol. 22, pp. 148-150, February 1997.
- [212] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*. Cambridge University Press, second ed., 1992.
- [213] J. M. Soto-Crespo, D. R. Heatley, and E. M. Wright, "Stability of the higher-bound states in a saturable self-focusing medium," *Physical Review A*, vol. 44, no. 1, 1991.
- [214] L. Allen and J. H. Eberly, *Optical resonance and two-level atoms*. Dover, 1987.
- [215] R. W. Boyd, *Nonlinear Optics*. Academic Press, 1992.
- [216] R. W. Equall, Y. Sun, R. L. Cone, and R. M. Macfarlane, "Ultraslow optical dephasing in $\text{Eu}^{3+}:\text{Y}_2\text{SiO}_5$," *Physical Review Letters*, vol. 72, pp. 2179-2182, April 1994.
- [217] T. W. Mossberg, "Time-domain frequency-selective optical data storage," *Optics Letters*, vol. 7, pp. 77-79, February 1982.
- [218] P. Saari, R. Kaarli, and A. Rebane, "Picosecond time- and space-domain holography by photochemical hole burning," *Journal of the Optical Society of America B*, vol. 3, pp. 527-534, April 1986.
- [219] A. Renn and U. P. Wild, "Spectral hole burning and hologram storage," *Applied Optics*, vol. 26, pp. 4040-4041, 1987.
- [220] P. N. Butcher and D. Cotter, *The Elements of Nonlinear Optics*. Cambridge University Press, 1990.
- [221] Y. R. Shen, *The Principles of Nonlinear Optics*. John Wiley and Sons, 1984.
- [222] R. W. Hellwarth, A. Owyong, and N. George, "Origin of the nonlinear refractive index of liquid CCl_4 ," *Physical Review A*, vol. 4, pp. 2342-2347, December 1971.
- [223] T. Brabec and F. Krausz, "Nonlinear optical pulse propagation in the single-cycle regime," *Physical Review Letters*, vol. 78, pp. 3282-3285, April 1997.
- [224] C. N. Ironside, "Ultra-fast all-optical switching," *Contemporary Physics*, vol. 34, no. 1, pp. 1-18, 1993.
- [225] A. Yariv and P. Yeh, *Optical Waves in Crystals*. John Wiley & Sons, 1984.
- [226] R. McLeod, *Spectral-Domain Analysis and Design of Three-Dimensional Optical Switching and Computing Systems*. PhD thesis, University of Colorado, 1995.
- [227] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations I*. Springer series in computational mathematics, Springer-Verlag, 1987.
- [228] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Transactions on Antennas and Propagation*, vol. AP-14, pp. 302-307, May 1966.
- [229] R. W. Ziolkowski and J. B. Judkins, "Applications of the nonlinear finite difference time domain (NL-FDTD) method to pulse propagation in nonlinear media: self-focusing the linear-nonlinear interfaces," *Radio Science*, vol. 28, pp. 901-911, September-October 1993.
- [230] D. U. von Rosenberg, *Methods for the Numerical Solution of Partial Differential Equations*. American Elsevier Publishing Company, Inc., 1969.
- [231] F. Tappert, "Numerical solutions of the Korteweg-de Vries equation and its generalizations by the split-step Fourier method," in *Lectures in Applied Mathematics*, vol. 15, pp. 215-216, American Mathematical Society, 1974.

- [232] J. A. Fleck, J. R. Morris, and M. D. Feit, "Time-dependent propagation of high energy laser beams through the atmosphere," *Applied Physics*, vol. 10, pp. 129–160, 1976.
- [233] M. D. Feit and J. J. A. Fleck, "Beam nonparaxiality, filament formation, and beam breakup in the self-focusing of optical beams," *Journal of the Optical Society of America B*, vol. 5, pp. 633–640, March 1988.
- [234] D. Yevick and M. Glasner, "Forward wide-angle light propagation in semiconductor rib waveguides," *Optics Letters*, vol. 15, pp. 174–176, February 1990.
- [235] J. Arabat and J. Etchepare, "Nonresonant fifth-order nonlinearities induced by ultrashort intense pulses," *Journal of the Optical Society of America B*, vol. 10, pp. 2377–2382, December 1993.
- [236] F. W. Sears, M. W. Zemansky, and H. D. Young, *University Physics*. Addison-Wesley, seventh ed., 1987.
- [237] A. Yariv, *Quantum Electronics*. John-Wiley & Sons, third ed., 1989.
- [238] A. Höök, "Influence of stimulated raman scattering on cross-phase modulation between waves in optical fibers," *Optics Letters*, vol. 17, pp. 115–117, January 1992.
- [239] A. R. Mickelson, *Guided Wave Optics*. Van Nostrand Reinhold, 1993.
- [240] D. Marcuse, *Theory of dielectric optical waveguides*. Academic Press, 1991.